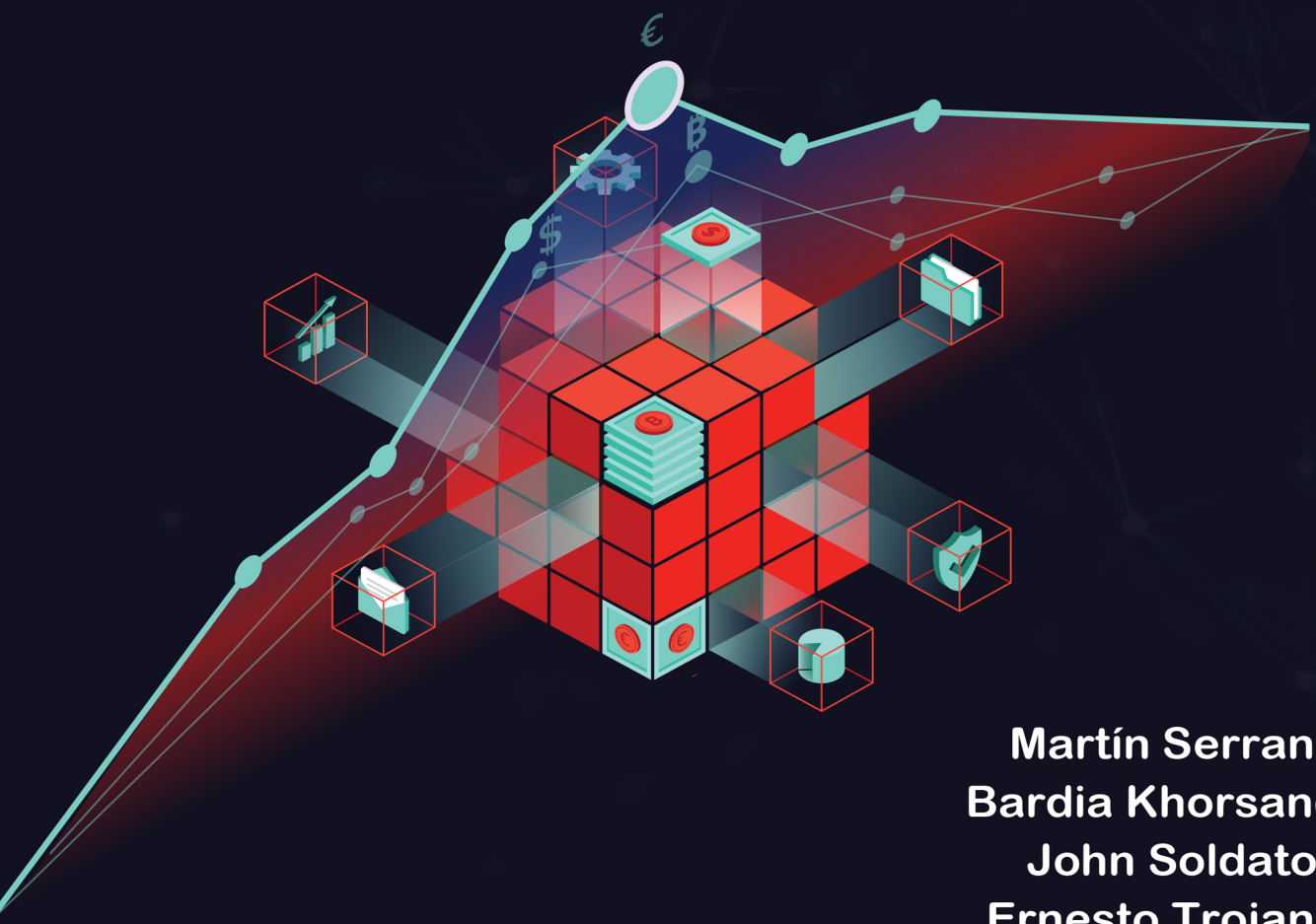
INFINITECH Book Series – Part 3

nowopen

Technical Financial Innovation, Solving the Interoperability Problems of Europe

The INFINITECH Legacy



Martín Serrano
Bardia Khorsand
John Soldatos
Ernesto Troiano
et al. (Editors)

Infinitech

now publishers

PART III:

Technical Financial Innovation, Solving the Interoperability Problems of Europe

The INFINITECH Legacy

Martín Serrano, Bardia Khorsand,
John Soldatos, Ernesto Troiano,
Juergen Neises, Pavlos Kranas, Kostis Perakis,
Alessandro Mamelli, Ignacio Elicegui,
Dimosthenis Kyriazis, George Makridis,
Gisela Sanchez and Marina Cugurra

now
the essence of knowledge

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

ISBN: 978-1-63828-232-7
E-ISBN: 978-1-63828-233-4
DOI: 10.1561/9781638282334

Copyright © 2023 Martín Serrano, Bardia Khorsand, John Soldatos, Ernesto Troiano, Juergen Neises, Pavlos Kranas, Kostis Perakis, Alessandro Mamelli, Ignacio Elicegui, Dimosthenis Kyriazis, George Makridis, Gisela Sanchez and Marina Cugurra

Suggested citation: Martín Serrano, Bardia Khorsand, John Soldatos, Ernesto Troiano, Juergen Neises, Pavlos Kranas, Kostis Perakis, Alessandro Mamelli, Ignacio Elicegui, Dimosthenis Kyriazis, George Makridis, Gisela Sanchez and Marina Cugurra. (2023). *INFINITECH Book Series – Part 3. Technical Financial Innovation, Solving the Interoperability Problems of Europe: The INFINTECH Legacy*. Boston–Delft: Now Publishers

The work will be available online open access and governed by the Creative Commons "Attribution-Non Commercial" License (CC BY-NC), according to https://creativecommons.org/licenses/by-nc/4.0/

Table of Contents

Preface ix

Who Should Read This Book? xi

What is Addressed in the Book Series? xiii

What is Covered in this INFINITECH Part III Book? xv

Acknowledgements xvii

Contributing Authors xix

Abstract xxix

Why FinTech xxxi

Chapter 1 INFINITECH Reference Architecture Overview 1

1.1 Reference Architecture for BigData/IOT/AI in Finance/Insurance . 1

1.2 INFINITECH Reference Architecture 2

Chapter 2 Innovative Technologies for the Financial Sector 7

2.1 Innovative Technologies for Financial Sector 7

2.2 INFINITECH Marketplace 13

2.2.1 Back-end 17

2.2.2 Front-end 18

VDIH [D8.5] 21

Training Activities 23

Courses 23

Workshops 23

Webinars 23

Innovation Services 23
Accelerator Programs 24
Social Login 26
Add New Information 27
Marketplace Usage Scenarios 29
Scenario 1: Upload a VDIH 30
Scenario 2: Consult a VDIH 30
2.3 Baseline Technologies 33
2.4 Interfaces (APIs) 33
2.4.1 APIs Related to Users 33
2.4.2 APIs Related to Descriptions 43
2.5 Search Functionality on Descriptions 65
2.6 APIs Related to Assets 67
2.7 Other Interfaces 72
2.8 Validation Scenarios 73
2.8.1 Authentication Scenario 73
2.8.2 Upload Description Scenario 74
2.8.3 Upload Asset Scenario 75
2.8.4 Retrieve Asset Scenario 75

Chapter 3 INFINITECH Technology Assets 79

3.1 Anti-Money Laundering 79
3.2 API 80
3.3 Applications 80
3.4 Blockchain 81
3.5 Data Marketplace 83
3.6 Database Management Systems 83
3.7 Dataset 84
3.8 Docker 86
3.8.1 Docker Containers 86
3.9 Finance 87
3.10 Financial Risk Estimation 88
3.11 Fraud Detection 89
3.12 Mobile Application 89
3.13 Model 90
3.14 Monitoring Tool 90
3.15 Notebook 90
3.16 Recommender 91
3.17 Screening Tool 92
3.18 Service 92

3.19 SMEs Business Risks 92
3.20 Software 92
3.21 TAHO 93
3.22 Visualization 93
3.23 Other 93

Chapter 4 INFINITECH Continuous Integration/Continuous Deployment Tools 95

4.1 Tools and Techniques for Testbeds and Sandboxes 95
4.1.1 Containers Benefits 96
4.1.2 Microservices Approach 97
4.1.3 Kubernetes Containers Orchestration 98
4.1.4 Kubernetes Architecture 99
4.2 INFINITECH Testbeds 101
4.3 INFINITECH Sandboxes 102

Chapter 5 INFINITECH Tools and Techniques for Management of Data Sets 105

5.1 Tools and Techniques for Management of Datasets 105
5.1.1 Data Sources and Data Access 105
5.1.2 Static Data Ingestion 107
5.1.3 Dynamic Data Ingestion 107
5.1.4 IoT Streaming 108
5.1.5 Direct Access to on Premise Data Sources 109
5.1.6 Blockchain Data Access 109
5.1.7 Third Party Data Access 110
5.1.8 Datasets Management from the Blueprint Reference Environment Perspective 110

Chapter 6 INFINITECH Reference Testbed and Sandbox 113

6.1 Overview of the INFINITECH Blueprint Reference Testbed 113
6.1.1 Development View 114
6.1.2 Creation of the EKS INFINITECH Cluster 117
6.1.3 Namespace, Network and Quote Policies 124
6.1.4 How to Recreate the Blueprint Testbed for a Specific INFINITECH Pilot 126
6.2 Blueprint Environment for Pilot 5b: Bank of Cyprus 127
6.2.1 Pilot Objectives 128
6.2.2 Pilot Workflow 128
6.2.3 Blueprint Reference Testbed Implementation 129

Chapter 7 Lessons Learnt from INFINITECH Pilots 133
7.1 Lessons Learnt [INFINITECH-Whitepaper-Lessons Learnt] 133
7.2 Cluster/Category 1: Smart, Reliable and Accurate Risk and Scoring Assessment 134
7.2.1 Pilot #2 – Real-time Risk Assessment in Investment Banking 134
7.2.2 Pilot #15 – Open Inter-banking Pilot 135
7.3 Cluster/Category 2: Personalized Retail and Investment Banking Services 135
7.3.1 Pilot #3 – Collaborative Customer-centric Data Analytics for Financial Services 136
7.3.2 Pilot #4 – Personalized Portfolio Management ("Why Private Banking cannot be for Everyone?") 138
7.3.3 Pilot #5b – Business Financial Management (BFM) Tools Delivering Smart Business Advice 139
7.3.4 Pilot #6 – Personalized Closed-Loop Investment Portfolio Management for Retail Customers 139
7.4 Cluster/Category 3: Predictive Financial Crime and Fraud Detection 140
7.4.1 Pilot #7 – Avoiding Financial Crime 140
7.4.2 Pilot #8 – Platform for Anti Money Laundering Supervision (PAMLS) 141
7.4.3 Pilot #9 – Analyzing Blockchain Transaction Graphs for Fraudulent Activities 142
7.4.4 Pilot #10 – Real-time Cybersecurity Analytics on Financial Transactions' BigData 143
7.4.5 Pilot #16 – Data Analytics Platform to Detect Payments Anomalies Linked to Money Laundering Events 144
7.5 Cluster/Category 4: Personalized Usage-Based Insurance (UBI) Pilots 145
7.5.1 Pilot #11 – Personalised Insurance Products Based on IoT Connected Vehicles 145
7.5.2 Pilot #12 – Real World Data for Novel Health-Insurance Products 147
7.6 Cluster/Category 5: Configurable and Personalized Insurance Products for SMEs and Agro-Insurance 149
7.6.1 PIlot #13 – Alternative/Automated Insurance Risk Selection – Product Recommendation for SME 150
7.6.2 Pilot #14 – Big Data and IoT for the Agricultural Insurance Industry 151

Chapter 8 Conclusions 153
8.1 Conclusions [D8.4] 153
Chapter 9 References 157
9.1 Infinitech Public Deliverables 157
9.2 References 161

This page intentionally left blank

Preface

The data is changing the way society and technology evolves, with the advent of IoT, Big Data, ML and AI, a rapid development in technology towards more human-centric applications has been envisaged. The finance and insurance sectors are not an exception and developments in FinTech and insurance-tech are in a phase of developing unique offerings.

It is very important to have a common understanding of the actual conditions in the financial and insurance sectors and how the technology can help to advance and evolve those conditions in a positive manner. By discussing the principles of the modern economy that make the modern financial sector and FinTech the most disruptive areas in today's global economy, a better understanding and knowledge will be acquired.

The use of data-driven approaches envisions many opportunities emerging for activating new channels of innovation on the local and global scale while at the same time catapulting opportunities for more disruptive human-centric services. Data-driven human-centric applications are at the same time the result of a shared vision from a natural evolution of technology and society. Experts in the financial and insurance sectors are looking at a dramatic change in how people think about global economy and at the same time the technology is facilitating the instruments for new ways of understanding, providing a common vision and identifying impacts in finance and insurance.

The INFINITECH book series is focused on addressing the need for clear information for better understanding of the foundations, principles and technologies for experts and non-technical experts that participate in the financial and insurance process and the constant need for innovation and new services across banks and insurance organizations.

The Editors and Contributors of this INFINITECH book series would like to thank the European Commission and the Science Foundation Ireland for their support in the planning and execution of the INFINTECH project that resulted in the preparation of this book. The recommendations and opinions, the provided and developed technologies alike experiences described in this book are those of the editors and contributors, and do not necessarily represent those of the European Commission or the Science Foundation Ireland.

Dr. J. Martin Serrano O.
INFINITECH Scientific Manager and Data Scientist
Adjunct Lecturer and Senior SFI Research Fellow at University of Galway
Data Science Institute – Insight SFI Research Centre for Data Analytics
Unit Head of Internet of Things, Stream Processing and Intelligent Systems Research Group
University of Galway, www.universityofgalway.ie | Ollscoil na Gaillimh
<jamiemartin.serranoorozco@universityofgalway.ie>
<martin.serrano@insight-centre.org>
<martin.serrano@nuigalway.ie>

Who Should Read This Book?

Financial & Insurance Regulators

The unique offering for non-technical experts but that participate in the financial regulatory process and of the core service to enable the sharing of innovation and new services across banks and insurance without exchanging any customer data.

General Public & Students

The power of understanding the future of FinTechs, their services and their ability to identify different methodologies indicators from a human perspective.

Entrepreneurs and SMEs

The most powerful tools to innovate, increase opportunities and increase the power of innovation into small and entrepreneurs to meet its full potential if there is good participation across the banking and insurance sector.

Technical Experts & Software Developers

The guide for technologies and legacy open and non-open sources as a guidebook for including the most recent experiences in Europe towards innovating technology for the financial and banking sectors.

This page intentionally left blank

What is Addressed in the Book Series?

"Concepts and Design Thinking Innovation addressing the Global Financial Needs"

In the first part of the INFINITECH book series we begin by discussing the principles of the modern economy that make the modern financial sector and FinTech the most disruptive areas in today's global economy. INFINITECH envision many opportunities emerging for activating new channels of innovation on the local and global scale while at the same time catapulting opportunities for more disruptive user- centric services. INFINITECH is at the same time the result of a shared vision from a representative global group of experts, providing a common vision and identifying impacts in the financial and insurance sectors.

"Methods and Design Principles for Financial Innovation, Explaining the Supply Side for Interoperability in Finance- and Insurance-Tech"

In the second part of the series we review the basic concepts for Fintech referring to the diversity in the use of technology to underpin the delivery of financial services. The demand and the supply side in the financial sector are demonstrated, and further discussed is why FinTech is the focus of industry nowadays and the meaning for waves of digitization. Financial technology (FinTech) and insurance technology (InsuranceTech) are rapidly transforming the financial and insurance services industry. We provide an overview of Reference Architecture (RA) for BigData, IoT and AI applications in the financial and insurance sectors (INFINITECH-RA). Moreover, this book reviews the concept of innovation and its application in INFINITECH, and innovative technologies provided by the project for financial sector practical examples.

"Technical Financial Innovation, Solving the Interoperability Problems of Europe"

The third book begins by providing a definition for FinTech as: The use of technology to underpin the delivery of financial services. This book further discusses why FinTech is the focus of industry nowadays as the waves of digitization and the way financial technology (FinTech) and insurance technology (InsuranceTech) are rapidly transforming the financial and insurance services industry. In this book technology assets that followed the Reference Architecture (RA) for BigData, IoT and AI applications are introduced. Moreover, the series of assets includes the domain area where applications from the INFINITECH innovation project and the concept of innovation for the financial sector are described. Further, we describe INFINITECH Marketplace and its components including details of available assets. Next, we provide descriptions of solutions developed in INFINITECH.

What is Covered in this INFINITECH Part III Book?

"Technical Financial Innovation, Solving the Interoperability Problems of Europe."

Technology frameworks and testbed tools (sandboxes) are popular these days. From the point of view of the deployment and testing, technology Pilots are defined in terms of resources used to deploy solutions (infrastructure) and Sandboxes (components). The different technologies developed within the context of the INFINITECH way initiative is considered as a bill of materials of the resources needed to perform demonstrators, proof of concepts and prototype solutions. The provided information can be used as input for configurators and cost structures to set up the testbeds and therefore it is extremely valuable to organizations from IT to financial and procurement departments.

The third book begins by providing a definition for FinTech as: the use of technology to underpin the delivery of financial services. This book further discusses why FinTech is the focus of industry nowadays as the waves of digitization and the way Financial Technology (FinTech) and Insurance Technology (InsuranceTech) are rapidly transforming the financial and insurance services industry. In this book technology assets that followed the Reference Architecture (RA) for BigData, IoT and AI applications are introduced. Moreover, the series of assets includes the domain area where applications from the INFINITECH innovation project and the concept of innovation for the financial sector are described. Further, we describe INFINITECH Marketplace and its components including details of available assets. Next, we provide descriptions of solutions developed in INFINITECH.

This page intentionally left blank

Acknowledgements

To our families for their incomparable affection, jollity and always understanding that scientific career is not a work but a lifestyle and encouraging us to be creative, and for their enormous patience during the time away from them, invested in our scientific endeavours and responsibilities and for their understanding about loving too much our professional life and its consequences, we love you!.

To all our friends and relatives for their comprehension when we have no time to spend with them and we are not able to join in time because we are in a conference or attending yet another meeting, for their attention and the interest they have been shown all this time to keep alive our friendship; be sure our sacrifices are well rewarded.

To all our colleagues, staff members and students at our respectively institutions, organisations and companies for patiently listening with apparent attention the descriptions and progress of our work and for the great experiences and the great time spent while working together with us and the contributions provided to culminate this book series project. In particular thanks by the support and confidence from all people that believed this series of books would be finished in time and also to those that didn't trusted on it, because thanks to them we were more motivated to culminate the project.

To the scientific community, this is our family when we are away and working far from our loved ones, for their incomparable affection, loyalty and always encouraging to be creative, and for their enormous patience during the time invested in understanding, presenting and providing feedback to new concepts and ideas, sincerely to you all, thanks a million!

Martín Serrano on Behalf of All Authors and Contributors

This page intentionally left blank

Contributing Authors

Achille Zappa
NUIG-Insight, Ireland

Adrien Besse
FTS, France

Aikaterini Karamargiou
NBG, Greece

Akshay Shetty
PRIVE, Germany

Alain Vailati
HPE, Italy

Alberto Crespo
ATOS, Spain

Alberto Danese
NEXI, Italy

Aleksandra Cargo
BOS, Slovenia

Alessandra Forlano
ENG, Italy

Alessandro Amicone
GFT, Italy

Alessandro Mamelli
HPE, Italy

Alessio Del Soldato
NEXI, Italy

Alex Acquier
NUIG-Insight, Ireland

Alexander Kostopoulos
RB, UK

Alper Sen
BOUN, Turkey

Andrea Becerra
CTAG, Spain

Andrea Grillo
PI, Greece

Andrea Toro
HPE, Italy

Andreas Politis
DYN, Greece

Angeliki Kitsiou
CP, Greece

Anja Rijavec Ursej
BOS, Slovenia

Ann Smith
BPFI, Ireland

Anna Semeniuk
PRIVE, Austria

Annalisa Ceccarelli
PI, Greece

Anne Elisabeth Lenel
ORT, France

Antonis Litke
INNOV, Cyprus

Antonis Skarpelis
CP, Greece

Ariana Polyviou
INNOV, Cyprus

Ariana Polyviou
INNOV, Cyprus

Aristodemos Pnevmatikakis
ISPRINT, Belgium

Baran Kilic
BOUN, Turkey

Barbara Cacciamani
ABILAB, Italy

Bardia Khorsand
NUIG-Insight, Ireland

Beatrice Paolone
PI, Greece

Bjoern Torkar
PRIVE, Germany

Borja Pintos Castro
GRAD, Spain

Brigitte Benerink
RRD, Netherland

Bruno Almeida
UNP, Portugal

Bruno Lepri
FBK, Italy

Can Ozturan
BOUN, Turkey

Candeago Candeago
FBK, Italy

Carlos Albo
WEA, Spain

Carmen Furquet
INSO, Spain

Carmen Perea
ATOS, Spain

Chi Hung Le
NUIG-Insight, Ireland

Christian Hanley
PRIVE, Austria

Christiane Grunloh
RRD, Netherland

Christina Katsikari
RB, UK

Christoforos Symvoulidis
SILO, Greece

Christopher Genillard
GEN, Germany

Chrysostomos Symvoulidis
UPRC, Greece

Claudia Amador
UNP, Portugal

Claudia Mertinger
FTSG, Germany

Craig Macdonald
GLA, UK

Cyril Armange
FI, France

Danae Lekka
ISPRINT, Belgium

Dario Francés
WEA, Spain

David Delgado
WEA, Spain

Davide Dalle Carbonare
ENG, Italy

Davide Profeta
ENG, Italy

Dejan Adamic
BOS, Slovenia

Diego Burgos
LXS, Spain

Dimitrios Kotios
UPRC, Greece

Dimitrios Miltiadou
UBI, Greece

Dimitris Drakoulis
INNOV, Cyprus

Dimitris Dres
INNOV, Cyprus

Dimosthenis Kyriazis
UPRC, Greece

Diogo Inácio
UNP, Portugal

Domenico Costantino
HPE, Italy

Domenico Messina
ENG, Italy

Dominik Hedderich
GEN, Germany

Dominique Faessel
ORT, France

Dustin Ciccardini
JRC, Germany

Ehsan Arefifar
CTAG, Spain

Elena Battistini
GFT, Italy

Elena Femenia
INSO, Spain

Eleni Mavrogalou
CP, Greece

Eleonora Ascolani
PI, Greece

Eoin Jordan
NUIG-Insight, Ireland

Erdem Oguz
AKTIF, Turkey

Ernesto Troiano
GFT, Italy

Eva Sotos Martinez
GRAD, Spain

Evelina Peristeri
SILO, Greece

Eymard Hooper
BOI, Ireland

Fabiana Fournier
IBM, Israel

Fabio Dezi
NEXI, Italy

Fabio Magrassi
GFT, Italy

Farid Meinkohn
ORT, France

Fethi Ata
AKTIF, Turkey

Filip Koprivec
JSI, Slovenia

Filipa Sousa
UNP, Portugal

Gabriele Gamberi
ABILAB, Italy

Gabriele Santin
FBK, Italy

Gary Thompson
BOI, Ireland

Gavin Purtill
BPFI, Ireland

George Fatouros
INNOV, Greece

George Giaglis
UNIC, Cyprus

George Karamanolis
CP, Greece

Georgios Makridis
UPRC, Greece

German Herrero
ATOS, Spain

Giacomo Toselli
SIA, Italy

Giancarlo Sfolcini
SIA, Italy

Giorgia Gazzarata
GFT, Italy

Giorgio Dabormida
GFT, Italy

Giorgio Roffo
GLA, UK

Giovanni Di Orio
NOVA, Portugal

Gisela Sanchez
FI, France

Giuseppe Avigliano
PI, Greece

Gokcehan Kara
BOUN, Turkey

Gregor Krzmanc
JSI, Slovenia

Gregor Zunic
JSI, Slovenia

Grigoris Mygdakos
AGRO, Greece

Guilherme De Brito
NOVA, Portugal

Harm op den Akker
ISPRINT, Belgium

Harm Opdenakker
RRD, Netherland

Hermie Hermens
RRD, Netherland

Hoan Nguyen
NUIG-Insight, Ireland

Iacopo De-angelis
PI, Greece

Iadh Ounis
GLA, UK

Iago Abad Fernandez
GRAD, Spain

Ian Godfrey
FTS, France

Ian Shiundu
RB, UK

Ignacio Elicegui
ATOS, Spain

Ilesh Dattani
ASSEN, Ireland

Ines Ortega-Fernandez
GRAD, Spain

Inna Skarbovsky
IBM, Israel

Irene Zattarin
GFT, Italy

Javier Rodriguez Viñas
GRAD, Spain

Javier Sanz-Cruzado Puig
GLA, UK

Javier Yepez Martínez
GRAD, Spain

Jelena Milosevic
BOS, Slovenia

John Soldatos
INNOV, Cyprus

Jonathan Gay
ASSEN, Ireland

Juan Mahilo
LXS, Spain

Juergen Neises
FTSG, Germany

Julian Schillinger
PRIVE, Germany

Julien Mousset
PRIVE, Germany

Klaudija Jurkosek-Seitl
BOS, Slovenia

Klaus Brisch
DWF, Germany

Klemen Kenda
JSI, Slovenia

Konstantina Kostopoulou
ISPRINT, Belgium

Konstantina Tripodi
JRC, Germany

Konstantina Zafeiri
NBG, Greece

Kostas Perakis
UBI, Greece

Lambis Dionysopoulos
UNIC, Cyprus

Lena Neidhardt
GEN, Germany

Lex Vanvelsen
RRD, Netherland

Lilian Adkinson Orellana
GRAD, Spain

Luca Latella
NEXI, Italy

Lucile Aniksztejn
FI, France

Lukas Linden
GEN, Germany

Maanasa Srikrishna
GLA, UK

Machi Simeonidou
AGRO, Greece

Mads Tingsgard
CPH, Denmark

Magdalena Schmid
GEN, Germany

Maja Skrjanc
JSI, Slovenia

Manolis Syllignakis
NBG, Greece

Manuela Masci
PI, Greece

Marc Meerkamp
DWF, Germany

Marcio Mateus
UNP, Portugal

Marco Avallone
PI, Greece

Marco Crabu
ABILAB, Italy

Marco Muller Terjung
DWF, Germany

Marco Pistore
FBK, Italy

Marco Rotoloni
ABILAB, Italy

Marco Spallaccini
PI, Greece

Marcos Alvarez Diaz
GRAD, Spain

Marcos Cabeza
CTAG, Spain

Margarita Khokhlova
FTS, France

Maria José Poveda
WEA, Spain

Maria de Vries
GEN, Germany

Maria Smyth
NUIG-Insight, Ireland

Marian Hurmuz
RRD, Netherland

Marianna Charalambous
UNIC, Cyprus

Mariarosaria Russo
ENG, Italy

Marina Cugurra
GFT, Italy

Marina Rodriguez Hidalgo
LIB, Spain

Marinos Xynarianos
CP, Greece

Mario Maawad Marcos
CXB, Spain

Mario Trinchera
ABILAB, Italy

Marko Grobelnik
JSI, Slovenia

Marta Sestelo
GRAD, Spain

Martin J. Serrano Orozco
NUIG-Insight, Ireland

Massimiliano Aschi
PI, Greece

Massimiliano Aschi
PI, Greece

Massimiliano Hocevar
PI, Greece

Matej Koletnik
BOS, Slovenia

Matteo Falsetta
GFT, Italy

Matteo Gerosa
FBK, Italy

Maurizio Ferraris
GFT, Italy

Maurizio Megliola
GFT, Italy

Maximilien Nayaradou
FI, France

Michael Concannon
BPFI, Ireland

Michael Michalakoukos
DYN, Greece

Michael Psalidas
CP, Greece

Misu Helal Ali
FTS, France

Mitja Jermol
JSI, Slovenia

Mojca Trstenjak
BOS, Slovenia

Nadia Roberti
PI, Greece

Napoleon Liontos
CP, Greece

Neil Giles
TAH, UK

Nial O'Brolchain
NUIG-Insight, Ireland

Niarchos Vasilios
NBG, Greece

Nicola Masi
ENG, Italy

Nikolaos Kapsoulis
INNOV, Cyprus

Nikos Drosos
SILO, Greece

Nikos Droukas
NBG, Greece

Nuria Ituarte Aranda
ATOS, Spain

Oliver Sjastedt
CPH, Denmark

Omerbora Zeybek
AKTIF, Turkey

Orkan Metin
AKTIF, Turkey

Pablo Carballo
PRIVE, Germany

Padraig Flannery
BOI, Ireland

Palmira Aldeguar
LIB, Spain

Paolo Testa
NEXI, Italy

Patrick Karlsson
RB, UK

Patrizio Sangermano
PI, Greece

Paul Lefrere
CCA, France

Pavlos Kranas
LXS, Spain

Pedro Malo
NOVA, Portugal

Perdikouri Eleni
NBG, Greece

Petra Ristau
JRC, Germany

Phil Atherton
TAH, UK

Prokopaki Georgia
NBG, Greece

Qaiser Mehmood
NUIG-Insight, Ireland

Raman Kazhamiakin
FBK, Italy

Ramon Martin de Pozuelo
CXB, Spain

Rebeca Jiménez
WEA, Spain

Rene Danzinger
PRIVE, Austria

Ricard Bruguera
WEA, Spain

Ricardo Jimenez-Peris
LXS, Spain

Richard McCreadie
GLA, UK

Richard Walsh
BPFI, Ireland

Rishabh Chandaliya
NUIG-Insight, Ireland

Roger Ferrandis
WEA, Spain

Roland Meier
PRIVE, Austria

Roman Benito
LIB, Spain

Sabina Podkriznik
BOS, Slovenia

Sara El Kortbi Martinez
GRAD, Spain

Saso Crnugelj
BOS, Slovenia

Silvio Walser
BOC, Cyprus

Simon Schou
CPH, Denmark

Simone Centellegher
FBK, Italy

Sofoklis Kyriazakos
ISPRINT, Belgium

Spyros Spanos
ISPRINT, Belgium

Stamatis Pitsios
UBI, Greece

Stathis Kanavos
ISPRINT, Belgium

Stefano Gatti
NEXI, Italy

Stelios Kotsopoulos
AGRO, Greece

Stelios Mantas
NBG, Greece

Stelios Pantelopoulos
SILO, Greece

Stephanie Jansen-kosterink
RRD, Netherland

Susanna Bonura
ENG, Italy

Susanna Bonura
ENG, Italy

Tanja Zdolsek draksler
JSI, Slovenia

Teoman Onat
PRIVE, Germany

Teresa Spada
ABILAB, Italy

Theodoros Kotzastavros
CP, Greece

Theodoros Arnaoutoglou
CP, Greece

Thomas Diesinger
GEN, Germany

Thomas Krogh
CPH, Denmark

Thorsten Jansen
DWF, Germany

Tiago Teixeira
UNP, Portugal

Vaia Gousdova
SILO, Greece

Vasilis Koukos
SILO, Greece

Vasilis Koukos
UPRC, Greece

Vicent Sebastiá
WEA, Spain

Vicky Foteinou
CP, Greece

Victoria Michailidou
RB, UK

Vito Morreale
ENG, Italy

Vittorio Monferrino
GFT, Italy

Yasar Khan
NUIG-Insight, Ireland

Ziga Bucaj
BOS, Slovenia

This page intentionally left blank

Abstract

Technology frameworks and testbed tools (sandboxes) are popular these days. From the point of view of the deployment and testing, technology Pilots are defined in terms of resources used to deploy solutions (infrastructure) and Sandboxes (components).

The different technologies developed within the context of the INFINITECH way initiative is considered as a bill of materials of the resources needed to perform demonstrators, proof of concepts and prototype solutions. The provided information can actually be used as input for configurators and cost structures to set up the testbeds and therefore it is extremely valuable to organizations from IT to financial and procurement departments.

The third book begins by providing a definition for FinTech as: *The use of technology to underpin the delivery of financial services*. This book further discusses why FinTech is the focus of industry nowadays as the waves of digitization and the way Financial Technology (FinTech) and Insurance Technology (InsuranceTech) are rapidly transforming the financial and insurance services industry.

In this book technology assets that followed the Reference Architecture (RA) for BigData, IoT and AI applications are introduced. Moreover, the series of assets includes the domain area where applications from the INFINITECH innovation project and the concept of innovation for the financial sector are described. Further, we describe INFINITECH Marketplace and its components including details of available assets. Next, we provide descriptions of solutions developed in INFINITECH.

This page intentionally left blank

Why FinTech

The term "FinTech" has been bandied around for a few years to describe the transformation taking place in the financial services markets in both the developed and the developing world. At the same time, forms of payment such as cheques are being displaced with cards. The cards are then being disrupted by contactless payment and digital wallets, which are now becoming the norm. Juniper defines fintech as: *The use of technology to underpin the delivery of financial services*. [The Future of FinTech – The New Standard]

The emergence of FinTech is increasingly blurring the lines between the two, previously very distinct, Financial Services (FS) and Technology (Tech), sectors. FinTech brings together the best of both these sectors to disrupt the financial services industry by enhancing customer experience, increasing the speed of service, and reducing operating cost through digitization. [deloitte uk human capital]

Historically, the financial and insurance services sectors have been quite resistant to technology disruption. However, this is no longer the case as the waves of digitization, Financial Technology (FinTech) and Insurance Technology (InsuranceTech) are rapidly transforming the financial and insurance services industry [Dietz16], [PwC17]. This is evident in the momentum and tangible growth of FinTech/InsuranceTech enterprises and in the volume of relevant investments: Over $23 billion of venture capital and growth equity has been allocated to FinTech innovations during 2011–2014, while $12.2 billion was deployed in 2014 alone. Moreover, a recent McKinsey & Co study revealed that the number of FinTech startups in 2016 exceeded 2.000, from approx. 800 in 2015. Furthermore, the vast majority of global banks and investment firms have already planned to increase their FinTech/InsuranceTech investments with a view to yielding a 20% average return on their investments. Beyond FinTech/InsuranceTech, financial institutions and insurance organizations are heavily investing in their digital transformation,

as a means of improving the efficiency of their business processes and optimizing their decision making.

The vast majority of digital transformation applications for the finance and insurance sectors are data intensive. This hold for applications in different areas such as retail banking, corporate banking, payments, investment banking, capital markets, insurance services, financial services security and more.

All these applications leverage very large datasets from legacy banking systems (e.g., customer accounts, customer transactions, investment portfolio data), which they combine with other data sources such as financial markets data, regulatory datasets, real-time retail transactions and more. With the advent of Internet-of-Things (IoT) devices and applications (e.g., fitbits, smart phones, smart home devices), several FinTech/InsuranceTech applications can take advantage of contextual data to offer better quality of service at a more competitive cost (e.g., personalized healthcare insurance based on medical devices and improved car insurance based on connected car sensors).

Furthermore, alternative data sources (e.g., social media and on-line news) provide opportunities for new more automated, personalized, and accurate services. Moreover, recent advances in data storage and processing technologies (including advances in Artificial Intelligence (AI) and blockchain technologies) provide new opportunities for exploiting the above-listed massive datasets and are expected to stimulate more investments in digital finance/insurance services.

Overall, financial and insurance organizations take advantage of BigData and IoT technologies to improve the accuracy and cost-effectiveness of their services, as well as the overall value that they provide to their customers. Nevertheless, despite early deployment instances, there are still many challenges that have to be overcome prior to leveraging the full potential of BigData/IoT/AI in the finance and insurance sectors, which could also act a catalyst for attracting more investments and for significantly improving the competitiveness of enterprises in these sectors.

In particular, financial institutions and insurance organizations are currently faced with the following challenges:

- Data Fragmentation and Interoperability Barriers.
- Limitations for Cost-Effective Real-Time Analytics.
- Regulatory Barriers.
- Data Availability Barriers.
- Lack of Blueprint Architectures for BigData Applications.
- No Validated Business Models.

In order to address these challenges, and leverage the full potential of BigData, IoT and AI and IoT in finance/insurance, there is a need for developments in several parallel streams, including:

- Technical/Technological Developments.
- Development of Experimentation Infrastructures (Testbeds & Sandboxes).
- Validation of Novel Business Models.

This page intentionally left blank

DOI: 10.1561/9781638282334.ch1

Chapter 1

INFINITECH Reference Architecture Overview

1.1 Reference Architecture for BigData/IOT/AI in Finance/Insurance

The aim of this book series is to review and validate the Reference Architecture (RA) for BigData, IoT and AI applications in the financial and insurance sectors (INFINITECH-RA), which will serve as a blueprint for the rapid and cost effective solutions development and deployment. The INFINITECH-RA will specify a set of building blocks that will support advanced BigData, AI and IoT applications. These building blocks will support scalable, unified and interoperable data collection from different sources and databases (e.g., OLTP-On-Line Transactional Processing-, OLAP-On-line Analytical Processing-, Data Lakes, SQL databases, NoSQL databases, alternative data sources), efficient real-time predictive analytics, multi-channel/Omni-channel interactions, data governance functionalities, as well as interoperable data sharing and interactions between stakeholders of the financial & insurance value chains. INFINITECH-RA will specify the structuring principles that will drive the integration of these building blocks in real-life solutions. The INFINITECH-RA will serve as a basis for designing, developing, and deploying novel BigData, AI and IoT solutions that feature "SHARP" (Smart, Holistic,

Autonomy, Personalized and Regulatory Compliance) characteristics. The project will also provide a number of blueprints for developing and deploying solutions aligned to the INFINITECH-RA. The blueprints will be based on the elaboration of different designs and deployment configurations that will be tailored to the needs of specific solutions. Both the INFINITECH-RA and its relevant blueprints will address functionalities that are prioritized as part of the SRIA (Strategic Research and Innovation Agenda) of the BDVA (BigData Value Association), while considering and consolidating concepts from RAs introduced by relevant standardization bodies and associations. [the proposal]

It will provide an overall of the technical requirements and specifications driving the project, as well as the detailed specification of the INFINITECH data models, technology/regulatory building blocks and the INFINITECH-RA. In here the following objectives are satisfied: (i) To articulate stakeholders' requirements regarding BigData and IoT-based services with SHARP properties in the financial and insurance sectors; (ii) To refine and detail the SHARP properties of various services in the target sectors; (iii) To analyze the background BigData and IoT platforms that will support the pilots and testbeds of the project, and to detail how they will be enhanced in order to empower the INFINITECH vision; (iv) To specify the security and regulatory compliance requirements of the INFINITECH services, while at the same time specifying the relevant solutions to be used in the project; (v) To specify the capabilities of the testbeds that will support the development and deployment of SHARP services; (vi) To specify the INFINITECH-RA.

1.2 INFINITECH Reference Architecture

Reference Architecture (RA) of the INFINITECH project aimed to develop Smart, Autonomous and Personalized Services in the European Finance and Insurance Services Ecosystem. The INFINITECH partners have selected a methodology to work on the RA, identifying it in the "4+1" architectural view model, which is presented in the document. The methodology is based on five different views, from which the structure of the system can be analyzed (logical view, process view, development view, physical view, and scenarios). Moreover, it will be demonstrated that all the functionalities of INFINITECH environment are properly covered by this model. [D2.13]

The State-of-the-Art survey underlines that some already existing Reference Architectures provide substantial input to INFINITECH, such as the pipelined and workflow approach to support the functionalities of the different Pilots and Use Cases of the project. Relevant inputs to the task have been considered, in particular the input coming from use-cases considered in task "User Stories and

Analysis of Stakeholders' Requirements" and a cross reference matrix. Finally, a layered and high-level reference view and a detailed logical view of the RA are presented. Different layers have been identified (infrastructure, data management and protection, data processing and architecture, analytics, interface, and presentation/visualization). The layers are mainly a mean of classification of the building blocks to form different workflows. The resulting RA provides a schema for building solid workflows and ensures full communication and interaction between all the building blocks, from the data source level (at the infrastructure level of the organizations) up to the Data Stores and Processing Analytics to presentation and visualization applications. High-Performance Computing (HPC) can be distributed at nodes within the platform supporting a high degree of scalability. [D2.12 to D2.15]

Moreover, RA considers for external data sources such as public and private Data Lakes, IoT networks and Blockchains. A list of identified building blocks provides the basic functionalities of the INFINITECH reference sandbox for a more general class of use cases. Building blocks are identified where existing technologies are available while other components and designed, implemented, and integrated during the INFINTEHC project duration as explained in public deliverables but summarized in this series book. [D2.13]

The validity of the RA has been proved by mapping the workflows of the pilots of the projects, ultimately proving the conceptual approach of the INFINITECH RA. The RA constitutes a living solution constantly verified during the continuous project development and with the different pilots. Moreover, the Consortium will promote the RA, along with its methodology and technological advancements, during project dissemination as a more general solution applicable to a broader set of different use cases beyond the original scope in the Financial and Insurance sectors whenever Big Data and AI are to be considered. [D2.13]

RAs are designed for facilitating design and developments of concrete technological architectures, mostly in the IT domain, reducing risks with proven components, all while improving overall communication within an organization. Real drawbacks and benefits of RAs have been analyzed with respect to the project's Pilots. RA facilitated development of concrete IT architectures and reduced maintenance costs. In general, the value of RAs can be summarized in the following points: [D2.15]

- Reduction of development and maintenance costs of systems.
- Facilitation of communication between important stakeholders.
- Reduction of risks.

Typically, when a system is designed without a RA, an organization may accumulate technical risks and end up with a complex and non-optimal implementation architecture. [D2.15]

In the industry, complex infrastructures for big data systems and high-performance computing (HPC) have been developed and proved to sustain intensive data processing services (Netflix, Facebook, Twitter, LinkedIn etc.). The architectures and technologies of world class infrastructure have been published and RAs have been designed and proposed. However, very few solutions have been published for the Financial and Insurance sectors and this paper aims to partially cover the gap. In the following sections, some relevant Reference Architectures and Models will be considered along with their relevance to the domain sector at which INFINITECH is aimed. [D2.15]

The purpose of a Reference Architecture is to provide a conceptual and logical schema for solutions to a large class of problems. In the INFINITECH project the domain is as vast as the Financial and Insurance Sectors where most of the applications are data-driven. The class of problems of the project (pilots and use cases) and in general the service management of financial institutions and insurance companies are largely based on data that should be managed in the safest and most protective way. [D2.13]

In these domains customers' enormous data sets must be processed to derive information with the purpose to provide better and more competitive services respecting the complex and sometimes conflicting regulatory frameworks such as privacy, security, interoperability, etc. [D2.13]

Therefore, a reference architecture should have explored in advance the specific domains in which the class of problems must find solutions providing a general model to which stakeholders (end-users, business owners, designers, data scientists, developers, maintainers etc.) can refer for best practices in the specific problem-solution space. In information technology, a RA can be used to check solutions to a particular problem in that class against the best practices and specific technologies. The INFINITECH-RA is no exception, and it is the result of the analysis of the significant number of use cases in the project's pilots, their requirements (users' stories) and constraints (regulatory, sector and technological) as well as the state-of-the art technologies and similar architectures. [D2.13]

It is important to state what is the RA in the INFINITECH project:

- A set of views for the Logical, Process, Development and Physical implementations.
- A set of common scenarios referring to generic use cases.
- A way to verify the use cases' scenarios and solutions.
- A way to speak the same language among stakeholders.
- A way to leverage solutions referring to best practices and building blocks.
- A way to verify if constraints in requirements, regulatory, technical and logical have been addressed properly.

It is also important to state what the INFINITECH RA is NOT:

- A ready-to-deploy technological IT framework.
- A rigid and unmutable set of connecting building blocks.
- A set of mandatory rules for development and integration.
- A manual for implementation and rollouts.
- A one-size-fits-all recipe for all business cases.

This page intentionally left blank

DOI: 10.1561/9781638282334.ch2

Chapter 2

Innovative Technologies for the Financial Sector

2.1 Innovative Technologies for Financial Sector

The INFINITECH project rely in the capabilities of its partner members to produce value that is exploitable beyond the lab or Proof of Concept of ideas, The Figure 2.1 shows the INFINITECH Innovation Roadmap where expertise from academia and research converges with industrial products and exploitation plans that uses design principles and reference architectures into creating a reference implementation for the financial and insurance sectors. The participation of stakeholders also complements the activity and brings value to transform ideas and innovate. [Innovation Readiness Assessment]

The data management functionalities of existing BigData and IoT platforms are not sufficient for fulfilling the needs of financial/insurance sectors applications, in terms of the need for providing integrated and unified access over the wide range of fragmented "siloed" data, the need for handling data streams and data at rest at the same time, as well as the need for supporting cost-effective execution of advanced data mining algorithms.

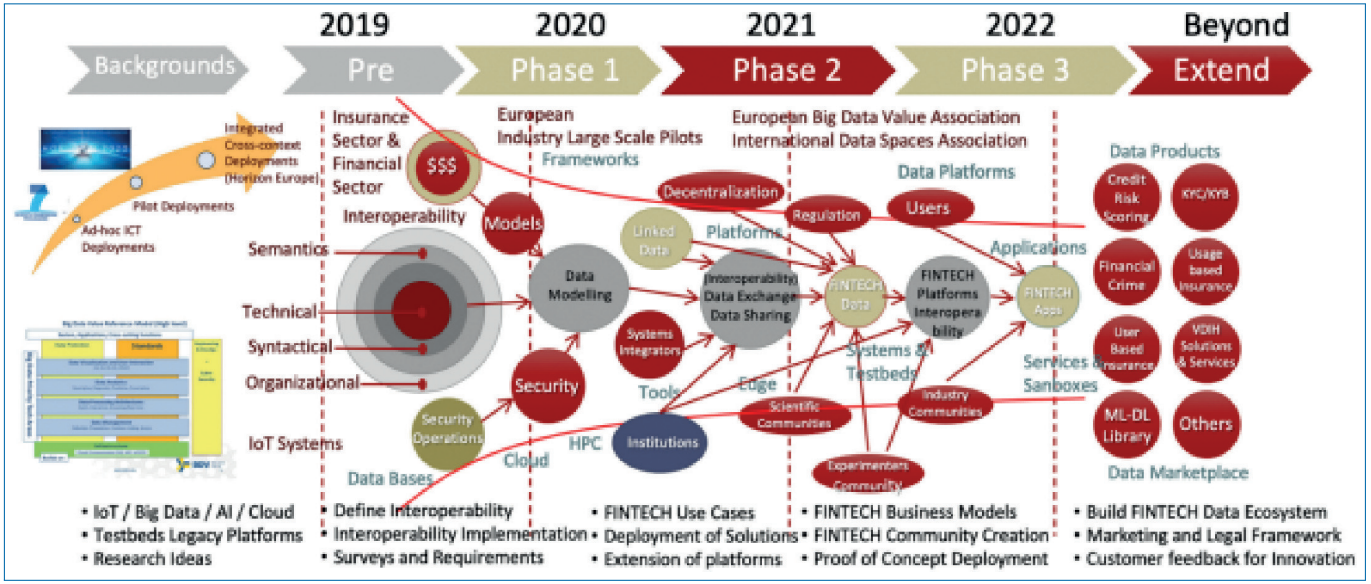


Figure 2.1 INFINITECH innovation roadmap.

INFINITECH will provide solutions for integrated data management over the wide range of databases and data sources used by BigData, IoT and AI applications in finance/insurance, including OLTP, OLAP, streaming data sources, structured data sources, unstructured data sources, semi-structured data source and more. The integrated data management solutions of the project will include: (i) A solution for integrated OLTP & OLAP data processing (i.e. so called HTAP1), which will facilitate unified access over data residing in both OLAP systems and operational databases without a need for tedious and expensive ETLs; (ii) A solution for integrated querying over both streaming data and data at rest; (iii) A solution for integrated and unified access over both SQL and NoSQL databases i.e. unified access over structured, unstructured and semi-structured data; and (iv) A solution for intelligent data pipelining and automated parallelization of stateful streaming engines to facilitate distributed high performance analytics.

INFINITECH will enable integrators and data scientists working on digital finance/insurance solutions to access data in a unified and integrated way, regardless of the underlying database/data store and the type of data handled (i.e. streaming/at-rest). This will facilitate the development of BigData/AI/IoT applications, through providing unified data access and query APIs, while at the same time improving the performance of queries and analytical algorithms.

Moreover, applications in the finance and insurance sectors must comply with many quite complex regulations. This holds for most BigData and IoT applications, which tend to be data-intensive and to involve complex data processing across multiple systems and stakeholders.

Project ensures that the INFINTECH platform is secure & regulatory compliant by design, through providing relevant data governance and regulatory compliance technologies in-line with the INFINITECH-RA. In particular, the project will design and implement technological building blocks for anonymization, authentication against eIDAS, management of consent, as well as data policies management. These building blocks will be correlated with the main regulations that have to be supported by financial and insurance organizations such as PSD2, 4MLD, MiFID II and GDPR. Regulatory compliance tools against these regulations using the project's tools will be developed. As part of the objective, BigData and IoT platforms/toolkits that will be used in the project will be enhanced with a data governance and regulatory compliance layer, which will be flexibly integrated in BigData and IoT applications in various configurations and in-line with their requirements and the INFINITECH-RA.

Further, existing BigData/IoT applications in the financial and insurance sectors form in most cases disaggregated (data) "silos", which are hardly interoperable with systems and application of other financial institutions and administrative domains. Likewise, there is very poor interoperability across the diverse datasets

that are typically collected and used in financial/insurance applications (including FinTech and InsuranceTech applications).

INFINITECH introduces building blocks for semantic interoperability and interoperable data exchange capabilities, as means of facilitating the development and deployment of innovative applications that span multiple systems and stakeholders in the financial supply chain (e.g., cross-border transactions, SWIFT payments, blockchain applications). As part of the objective, two solutions for interoperability and data exchange will be developed, including: (A) A centralized solution that will be based on the development of an interoperability database/registry supporting linking of diverse systems and datasets based on shared semantics, as well as semantically interoperable analytics. The semantic interoperability solution will leverage mainstream ontologies that are used in the financial sector (e.g., FIBO (Financial Industry Business Ontology) and FIGI (Financial Instrument Global Identifier)), along with distributed engines for massively parallelized and high-performance analytics over semantic data streams and related ontologies (i.e. high-performance execution of SPARQL queries). (B) A decentralized solution that will be based on the deployment of distributed ledger technologies and the use of permissioned blockchain concepts for decentralized trust. It will be based on infrastructures for permissioned blockchains that will become part of the INFINITECH solutions & testbeds to facilitate financial/insurance processes that involve cross-organization data exchange (e.g., credit risk scoring and KYC/KYB processes). The blockchain solution of the project will be augmented with functionalities for tokenization and secure/private-friendly querying. Tokenization will enable assets trading, while secure querying will enable querying blockchain data (e.g., customer data) without decrypting the source data as a means of providing strong privacy and data protection for applications (e.g., KYC, customer personalization) that need it.

Novel AI & BigData analytics applications for the finance and insurance sectors must combine a multitude of advanced analytics techniques (e.g., supervised, unsupervised and reinforcement learning models) over a great variety of data stores. To this end, they can greatly benefit from mechanisms and tools that facilitate access to analytics functions (including APIs), as well as from low-latency techniques & algorithms that could boost real-time analytics functions.

INFINITECH will develop a range of enablers for efficient, high-performance analytics that combine data from multiple sources and enable low latency, near real-time operations. To facilitate high-performance analytics INFINITECH will parallelize incremental algorithms that are commonly used in finance/insurance applications (e.g., clustering, collaborative filtering and frequent pattern matching) as a means of accelerating their execution. At the same time, to facilitate development of analytics applications, INFINITECH will provide a declarative analytics

framework that will enable analytics over diverse data sources based on conventional SQL-like primitives and in a way that handles the underlying complexity of data sources.

Moreover, the project will offer library of ML/DL algorithms for analytics in the finance/insurance sector, including both conventional algorithms and parallelized incremental algorithms. The project's ML/DL algorithms will also be made available as part of open data science frameworks and repositories such as OpenML, as a means of boosting their sustainability and wider use.

Overall, we can categorize certain challenges in the Fin/InsurTech sector as below:

- Unable to perform analytics over operational data.
  - Cannot provide real-time Business Intelligence.
- no integrated and unified access over fragmented "siloed" data.
- cannot perform real-time ingestion and analysis of real-time streams.
- no integrated query processing of data at-rest with in-flight.
- no cost-effective execution of advanced data mining algorithms.
  - data rates are dynamic, we must scale data and computation capabilities.
- many and quite complex regulations in finance/insurance sector.
- BigData/IoT applications tend to be data-intensive and to involve complex data processing across multiple systems and stakeholders.

To tackle these challenges INFINITECH developed certain technologies:

- Data management layer
  - HTAP (Hybrid Transactional/Analytics Processing) capabilities
  - Data ingestion at very high rate
  - Polyglot query processing
- Streaming processing framework
  - Integrated query processing of data *in-flight* and *at-rest*
  - Auto-scalability of parallelized streaming operators
  - Intelligent data pipelines
- Data regulation and governance
  - Provision of data governance mechanisms
  - Regulatory constraints tools

INFINITECH achievement can be grouped into the following categories:

- Enable data ingestion at very high rates with linear scalability, ideally used for data offloading.

- Provide data analytics over operational data (HTAP) while data is being ingested concurrently.
- Analytics on operational data, reduce the need for ETL, enabling real-time business intelligence.
- Support advanced analytics combining real-time streaming data with historical data.
- Implementation of change data capture mechanism via the Intelligent Data Pipelines.
- Unified framework for query processing over streaming/static data.
- FinFlik enables linear scalability of streaming operators, enabling their fully parallelization.
- Efficient resource consumption via the automation of scaling actions provided by the FinFlink library.
- Data Governance and Anonymization Framework.
- Policy Rules Orchestration Enforcement Framework.

The impact of these technologies applied can be listed as below:

- Analytics on operational data, reduce the need for ETL, enabling real-time business intelligence.
- Support advanced analytics combining real-time streaming data with historical data.
- Efficient resource consumption via the automation of scaling actions provided by the FinFlink library.
- Provide a good safeguard against data misuse, enabling regulatory compliance such as the GDPR.
- Set of privacy and utility metrics to better address privacy risk, preserving the utility of the data.
- Creation of user identities without sharing biometric information.
- Facilitates the orchestration of technologies for preserving privacy, data protection and security.
- Enables future compliance with new or changed or freshly-identified regulations.

To deliver on the assigned tasks, INFINITECH applied the following technologies:

1. A **pseudonymization tool** that automatically determines the best anonymization configuration for each application.
2. A **tool for anonymizing data** that automatically determines the best anonymization configuration for each application.

3. A **mobile digital user onboarding services** with virtual eID derived from government issued documents (ePassport or eID card).
4. A **solution for authenticating** citizens and businesses **against the eIDAS infrastructure**, providing a cross-border strong authentication mechanism based on eIDs.

2.2 INFINITECH Marketplace

The project's validated solutions at the technological or business level are made available at the project multi-sided market platform (marketplace) and/or a Virtualized Digital Innovation Hub (VDIH) for wider use and commercial exploitation. The market platform and the VDIH of the project are the points where the project will interact with other stakeholders of the digital finance/insurance and FinTech/InsuranceTech ecosystem. INFINITECH will build a community around its results i.e. based on stakeholders from all EU-28 countries that will engage in the market platform and/or the VDIH.

The INFINITECH multi-sided market platform offers BigData, AI, IoT, Blockchain and VDIH solutions is a public web-based environment with various APIs, able to store several types of assets that may derive/result from the separate procedures and mechanisms that are either implemented in the scope of the project or not (e.g. third-party contributions through hackathons/webinars etc.).

The market platform will integrate ready-to-use solutions and assets of the project, such as (synthetic) datasets/data assets, ML/DL algorithms, as well as validated turnkey solutions/applications for finance and insurance. At the same time, the innovation labs and FinTech/InsuranceTech clusters of the consortium will federate resources towards establishing a virtualized DIH, which will provide a host of innovation management services to incumbent financial/insurance organizations, but mainly to FinTech/InsuranceTech innovators. The project's market platform and VDIH will provide a single pan European entry point for accessing resources for BigData, IoT and AI innovations in the finance/insurance sectors. The market platform and the VDIH are the main enablers of the project's exploitation strategy.

INFINITECH will specify and implement the interactions of these building blocks, to support all stakeholders across the entire lifecycle of a novel BigData/IoT product or service for finance or insurance. This lifecycle includes several stages from the inception of the service to its technical and business validation. The latter validation is a prerequisite for the production use of the service, but also for making resources associated with the service available as part of the market platform and the VDIH of the project.

The multi-sided marketplace enables the utilization of Big Data and AI technologies, serving as a single end-point for the available assets found in the marketplace and includes algorithms, datasets, frameworks, webinars and lectures, docker containers and combined solutions, scalable and adaptable to fit into the business needs of companies and organizations, especially of the Fintech and InsuranceTech subsectors. The assets accommodated in the marketplace have been offered by various INFINITECH members, developers and data scientists and can be leveraged internally, increasing collaboration and knowledge transfer among INFINITECH, or externally by developers, service providers, tech companies and organization, all while promoting the results of the Project and the impact made to the Finance and the Insurance sectors. [D8.3]

INFINITECH provides a completely integrated environment enabling the utilization of big data and AI techniques in the finance and insurance sectors. The latter has become feasible through a set of technologies that enable exploitation of various datasets (obtained from different sources), optimized data management for these datasets (e.g., across diverse data stores), analytics with innovative algorithms covering a wide set of scenarios in the finance and insurance sectors, as well as use of tailored sandboxes on the underlying infrastructure layer for the execution of the aforementioned algorithms. Additionally, the development progress of INFINITECH pilots and provided sandboxes and testbeds have increased the INFINITECH ecosystem's resources, ranging from ML models and AI algorithms to IoT applications, Blockchain and a variety of ready-to-use solutions. [D8.3]

The INFINITECH solution goes beyond the utilization of analytics on specific datasets for a few pilots/use cases, by aiming at a generalized approach that will facilitate the exploitation of various analytics algorithms (provided both by INFINITECH researchers/partners and by 3rd party data analysts) on top of different datasets. To this end, the analytics algorithms need to be made available, to be described in terms of functionality, parameters, and offerings, to be accompanied with datasets that can be used by interested parties in order to validate their applicability and performance and to be offered as ready-to-be-executed solutions (e.g., containerized) in order to increase their utilization. [D8.3]

All of the above are representative functionalities of the INFINITECH market platform, which is at an operational stage (accessible at https://marketplace.infinitech-h2020.eu/) with various updates and functionalities introduced after the initial version. The marketplace holds and offers solutions for realizing big data and AI techniques in the finance and insurance sectors. [D8.3]

Based on the above, the INFINITECH's multi-sided market platform aims at being one of the project's main ambassadors to the big data and AI communities. It is a single, public and hybrid system with many different APIs, covering all the

different required perspectives of the platform. The various APIs developed related to users, descriptions and assets are described in this document. [D8.3]

The market platform offers big data and AI solutions, as well as IoT and Blockchain solutions, and VDIH Services. Thus, the INFINITECH market platform is a four-perspective, unified environment being able to store several types of assets (e.g., algorithms, descriptions of algorithms, evaluation and validation results, datasets, experimentation outcomes, etc.) in any format. [D8.3]

The INFINITECH marketplace is already deployed and available at https://marketplace.infinitech-h2020.eu/. It is currently populated with various assets and offerings which include: [D8.3]

- Algorithms
- Datasets
- Frameworks & mechanisms
- Scientific studies/tutorials
- Videos/webinars/lectures
- Experimentation results
- VDIH services
- End-to-end solutions
- Docker containers
- Combined solutions

The assets above have been provided by an extended audience, which includes INFINITECH Project members, developers, data scientists, VDIH services providers, as well as other authorized third-party users.

The available offerings could be leveraged by a variety of end-users, with increased potential and emphasis being given to FinTech and InsuranceTech subsectors. [D8.3]

With the progress of INFINITECH Project and included pilots' development, new scalable and adaptable assets are introduced on a regular basis and are made available to the end-users, which can then be deployed and executed based on their needs. Moreover, additional external stakeholders, organizations and service providers are also encouraged to accommodate their assets in the INFINITEH marketplace. [D8.3] The marketplace platform provides several functionalities that are mapped to different layers. In more detail, the back-end includes three layers (i.e. Assets Storage Layer, Assets Management Layer, and Interaction Layer), while the front-end includes one layer (i.e. Presentation Layer). The four layers of the marketplace along the primary functionalities are depicted in the Figure 2.2 below: [D8.3]

- The Assets Storage Layer (part of the back-end) is the layer in which the platform's offered assets are stored (See Figure 2.3).

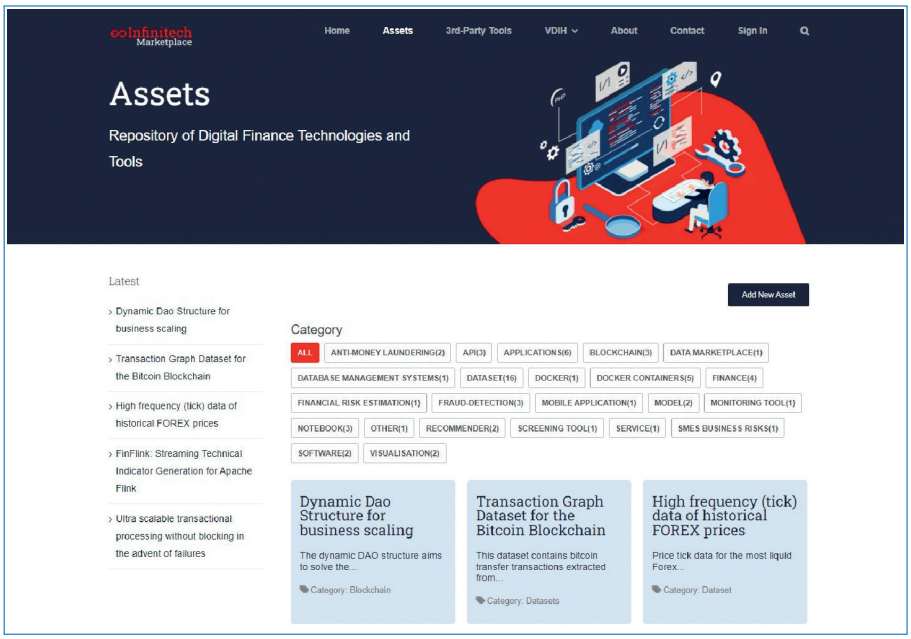


Figure 2.2. Snapshot from the "Assets" page on INFINITECH's Marketplace web page.

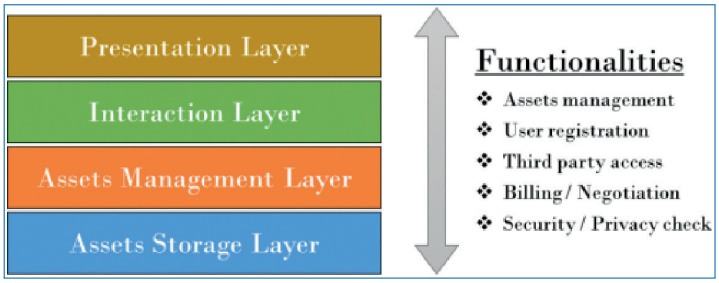


Figure 2.3. Market platform's layers and main functionalities.

- The Assets Management Layer (part of the back-end) delivers all the needed principles and techniques for the management of the Marketplace's assets.
- The Interaction Layer (part of the back-end) supports the communication between the market platform and its users (i.e. human users, and machine users), by providing discrete APIs for exploiting each different type of asset.
- The Presentation Layer(part of the front-end) provides the User Interface towards the different types of users that are willing to use the platform.

The market platform is structured around two core components, the back-end and the front-end. This approach contributes towards the platform's enhancements in terms of functionality as well as provides additional information and capabilities.

In that direction, the end-users are able to interact with the market platform through the front-end (through the presentation layer) that utilizes a user-friendly UI, while other additional services (e.g. from 3rd parties) can be implemented directly with the back-end (through the interaction layer). [D8.3]

The back-end, which contains structured information and the assets offered by the market platform, is considered the core of the marketplace and its functionalities. The front-end component is a user-friendly UI which presents to the users the offered content (the assets and their information), allowing them to interact with the platform in an easy and effective way. This section provides a short overview of the core components of the marketplace, describing some of their key features. [D8.3]

2.2.1 Back-end

The back-end is the main component of the marketplace. It consists of three different layers and implements the main functionalities for the assets management. The three levels are briefly described below. [D8.3]

The Assets Storage Layer is responsible for storing the assets that will be offered by the market platform. An essential component of this layer is the database that can store files in any format as well as additional information about the files provided. In this context, the type of database that is used is a document-oriented NoSQL database, which stores both JSON-like documents (the format of the descriptions files that are analyzed in the Assets Management Layer) and binary files, using extended specifications (e.g. file system). [D8.3]

The Assets Management Layer is responsible for the entire life cycle of the assets within the platform and offers all the principles and techniques for their management. Specifically, this layer handles the assets from the moment they are added to the platform through the APIs and then stored in the database (Assets Storage Layer), until they are to be deleted for any purpose from the platform. Through this layer, the market platform supports the CRUD operations and searching functionality, which are triggered by the corresponding APIs of the back-end (Assets Interaction Layer). The back-end is a REST API and receives different HTTP requests in order to perform an operation/ trigger a functionality. Moreover, there are mandatory description files for all available assets that contain metadata about the described asset (in JSON format). These description files are mandatory to make the assets searchable and retrievable by the end-users of the marketpalce. [D8.3]

The last layer, the Assets Interaction Layer, is responsible for supporting the communication between the market platform and its end-users. It implements the interfaces (APIs) of the back-end that will handle the back-end's operations. As

described before, these APIs receive HTTP requests that trigger the CRUD operations for both assets and description files. [D8.3]

2.2.2 Front-end

The front-end is the fourth layer of the market platform. It is a web-based server that presents the offered assets to the users, with a friendly UI. The front-end converts all interfaces of the back-end (REST API) into user friendly interfaces and provides automated forms and processes that make it easier for users to interact with the back-end and benefit from its stored assets. Therefore, it acts as an intermediate among the marketplace users and the back-end, sending the respective HTTP requests to the latter and presents its responses. [D8.3]

In short, the front-end allows users to register and log-in to the marketplace (user-based platform), upload their offered assets by filling out appropriate forms whose fields will be the content of the description files of the assets; search for assets according to various fields (title, asset's type, fields of use, provider, other metadata, etc.) that can be further filtered or even sorted by the number of views or the date they were uploaded to the marketplace, etc. Also, there is a page that presents in detail the information of the assets, and through this page, the users can retrieve the real assets, the files (See Figure 2.4). [D8.3]

There are certain updates and improvements to the INFINITECH Marketplace frontend. The structure initially defined was reorganized to accommodate the new content and facilitate the navigation within the INFINITECH Marketplace. New pages were created, and some existing ones were improved. This update also brought new functionalities to the INFINITECH Marketplace users.

The structure of the INFINITECH Marketplace was updated, the main update was in the VDIH page, previously named INFINITECH Academy. The name was changed to match the focus of the INFINITECH Marketplace that is "…establish a market platform that will provide access to the project's solutions, along with a Virtualized Digital Innovation Hub (VDIH) that will support innovators (FinTech/InsuranceTech) in their BigData/AI/IoT endeavors". In the Figure 2.5 it is possible to see the current structure of the INFINITECH Marketplace.

The diversity among VDIH resourcesled to the organization ofthe page into Training Activities and Innovation Services. While Training Activities page includes courses, webinars and workshops, the Innovation Services page include the accelerator programmes. For each type of VDIH content was created a new page, which brings together all the resources present on the platform related to that content.

The updates aren't restricted to the structure of the INFINITECH Marketplace, it was updated pages to support new functionalities and new pages was created to present new content.

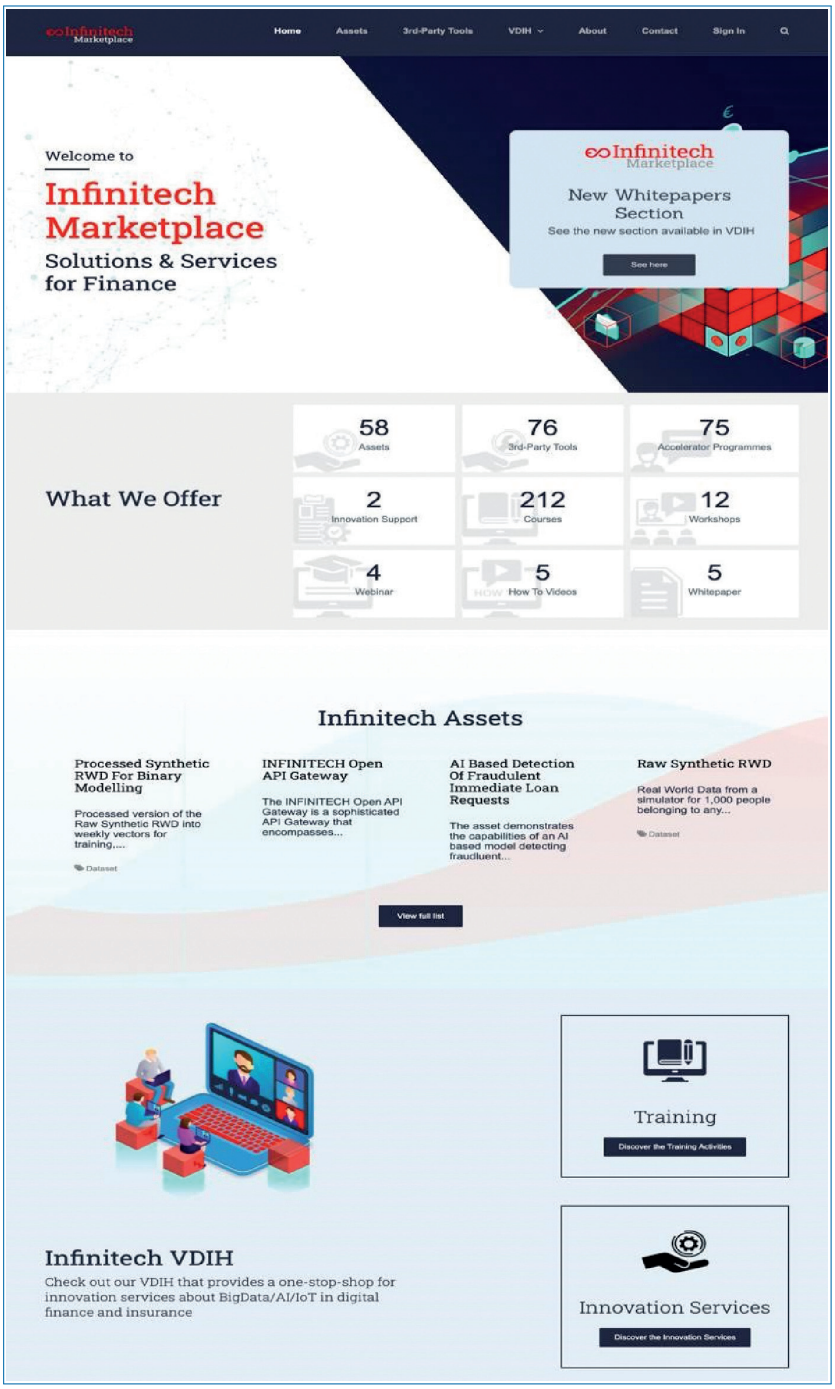


Figure 2.4. Snapshot of the INFINITECH marketplace home web page.

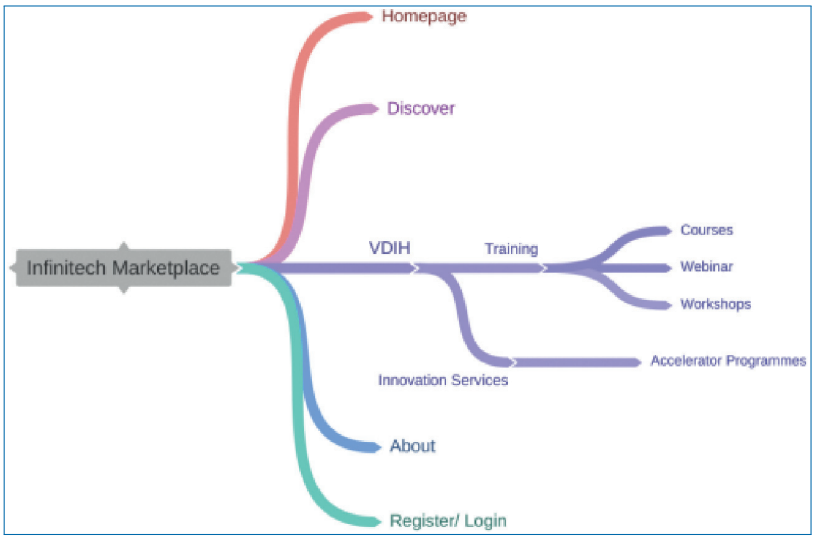


Figure 2.5. INFINITECH marketplace structure.

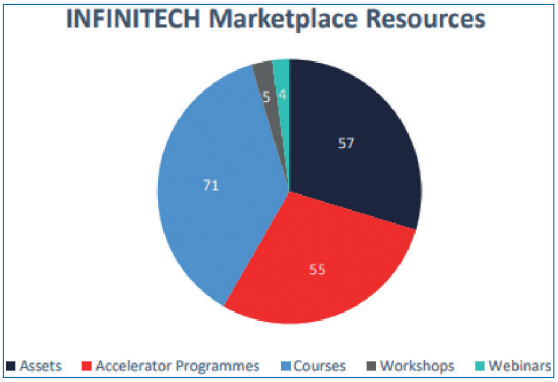


Figure 2.6. INFINITECH marketplace resources.

The statistics show the type and the number of resources available on the platform. As it can be seen in the Figure 2.6 above, presently the INFINITECH Marketplace contain assets, courses, workshops and accelerator programmes, corresponding to a total of 192 resources.

The INFINITECH Assets section provides access to the assets page and the INFINITECH VDIH section provides access to Training Activities and Innovation Services. Both sections were visually updated, making it easier to access the resources.

The last section of the Homepage is dedicated to new functionalities and has been updated to highlight social login (Sign in) and forms (Add New Information). The forms allow the INFINITECH users to provide information to the platform.

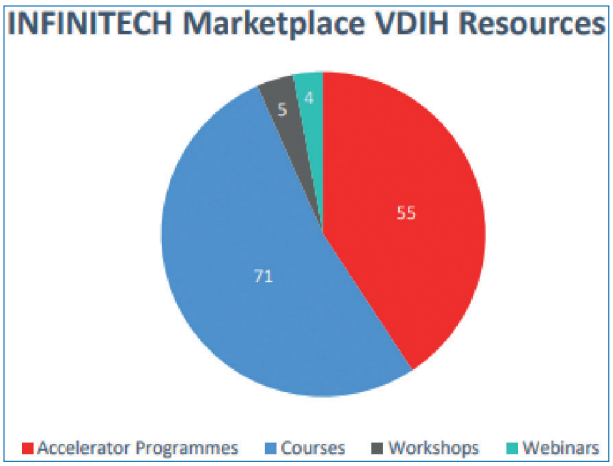


Figure 2.7. INFINITECH marketplace VDIH resources.

This new feature will be explained in the New Functionalities of Marketplace section.

VDIH [D8.5]

The VDIH name has been changed to match the focus of the INFINITECH Marketplace. The page also was organized into Training Activities and Innovation Services, while the Training Activities include courses, workshops and webinars, the Innovation Services includes acceleration programs. Accelerator programs are services to support and guide startups and SMEs on the creation of innovation through experts advisory, co-working space, education and skills development, among others.

As presented in Figure 2.7 above, the INFINITECH Marketplace currently has a total of 135 VDIH resources, which correspond to courses, workshops, webinar and accelerator programs.

All VDIH content pages are similar and Figure 2.8 shows the Courses page as a demonstration of these pages. The VDIH pages provide all the resources available on the platform related to that type of content, as well as provide some features for the user:

1. Filter by source: It allows to differentiate the source of the resources, if it belongs to INFINITECH partners or external entities, through the options: ALL, INFINITECH Course (in the case of the demonstration) or External.
2. Red Flag with the INFINITECH symbol: Related with the filter, it allows to differentiate the source of the resources, highlighting the resources which belongs to INFINITECH partners.

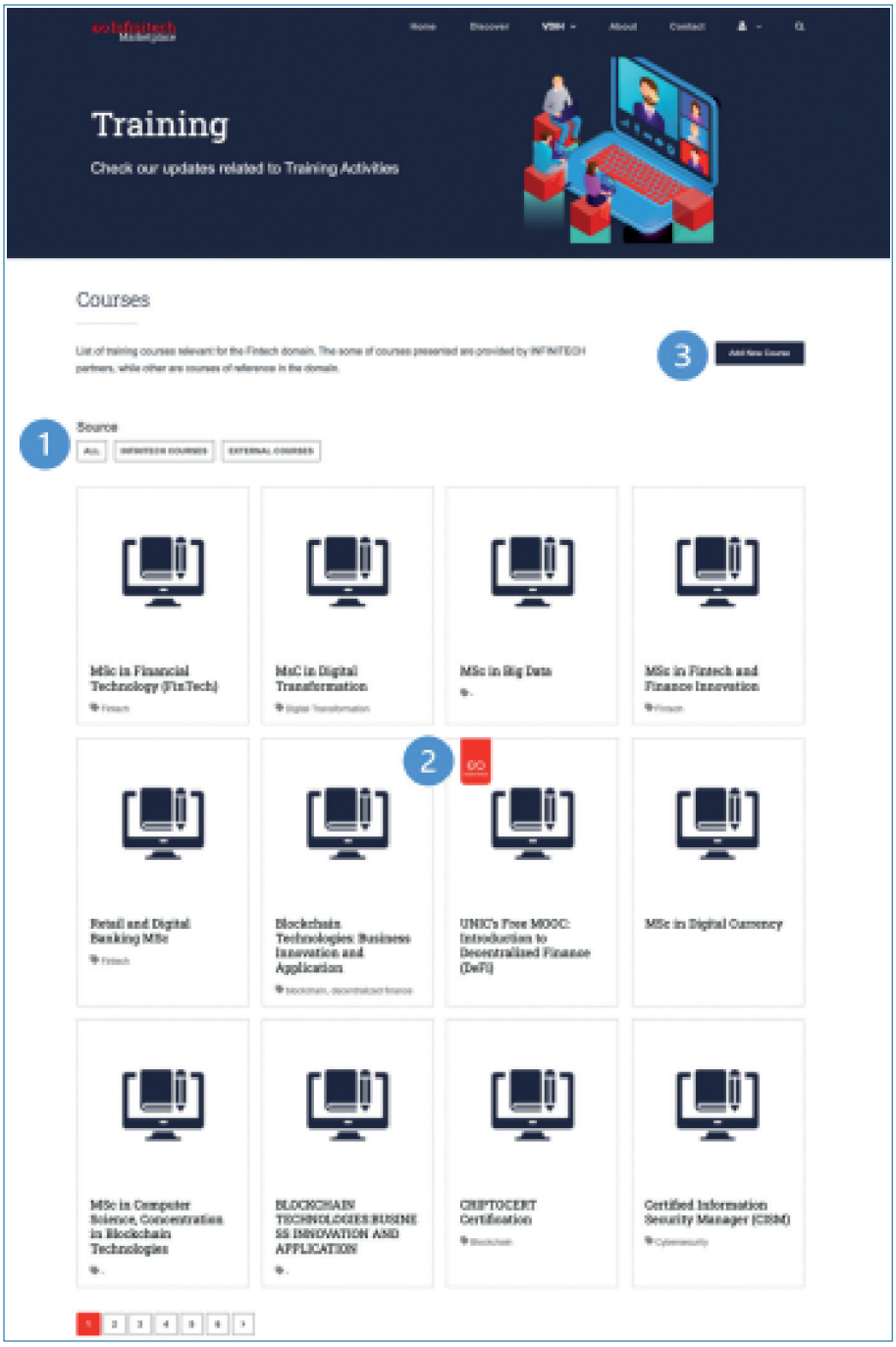


Figure 2.8. Courses page demonstrating features.

3. Add new resource: In this case, appear "Add new course", the button gives access to the course form. Each VDIH as its own form that allows the INFINITECH users to provide information. This feature will be explained in the New Functionalities of Marketplace section.

Training Activities

The INFINITECH Marketplace is a network with relevant resources about IoT, Blockchain, BigData and AI for finance/insurance. Training Activities is a new page of this network with a variety of courses, workshops and webinars that offer INFINITECH users the opportunity to improve their skills, knowledge or expertise.

Courses

The Courses page list all the courses available on the INFINITECH Marketplace. Each course has its own page, with information that characterizes it, such a brief description, who provided the course, difficult level, duration, cost, platform where the course will be available, among others. In addition to, the user also can know more about the course through the website present on the page.

Figures 2.9, 2.10 and 2.11 shows all courses available on the INFINITECH Marketplace so far, a total of 71 VDIH resources.

Workshops

The workshops page list all the workshops present on the INFINITECH Marketplace. Each workshop has its own page, with information that characterizes it, such a brief description, duration, who are the speakers, presentation files, videos and more like it can be viewed in Figure 2.12.

Webinars

The Webinars page brings together all the webinars available on the INFINITECH Marketplace. Each webinar has its own page, with some information that characterizes it, such as a brief description, date, organizer, as well as additional files and videos (See Figure 2.13).

Innovation Services

Innovation Services page was added to the INFINITECH Marketplace network and that's where we can find accelerator programs information. The Innovation Services intends to be a page where the user can find content to get support to a self-determined and sustainable approach to digital pioneering and help to accelerate businesses and guide it into the future to achieve better results.

Courses	Link
MSc in Blockchain and Digital Currency	link
UNIC's Free MOOC: Introduction to Digital Currencies	link
MSc in Computer Science, Concentration in Blockchain Technologies	link
Digital Transformation in Financial Services Specialization	link
Innovation Strategy: Developing Your Fintech strategy	link
Future Development in Supply Chain Finance and Blockchain Technology	link
Machine Learning and Reinforcement Learning in Finance Specialization	link
FinTech: Finance Industry Transformation and Regulation Specialization	link
Using Machine Learning in Trading and Finance	link
Reinforcement Learning in Finance	link
Fundamentals of Machine Learning in Finance	link
Overview of Advanced Methods of Reinforcement Learning in Finance	link
Supply Chain Finance and Blockchain Technology Specialization	link
Blockchain, Cryptoassets, and Decentralized Finance	link
Introduction to Blockchain for Financial Services	link
Blockchain Transformations of Financial Services	link
Blockchain Revolution Specialization	link
Digital transformation-Finance-Strategies	link
Open Banking, PSD2 and GDPR. FinTech	link
Finance and Accounting in the Digital Age	link
Artificial Intelligence for Finance, Accounting & Auditing	link
Complete 2-in-1 Python for Business and Finance Bootcamp	link
Practical Machine Learning: Real World Projects In Finance	link
AI for Finance	link
Mastering FinTech and Machine Learning!	link
Machine Learning Practical: Real World Projects In Finance	link
FinTech – Prepare for the revolution in Finance	link
Blockchain For Business+Finance Professionals 2019 – NEW	link

Figure 2.9. Courses available on the INFINITECH marketplace.

Accelerator Programs

The Accelerator Programs page brings together all the accelerator programs existent on the INFINITECH Marketplace. Each accelerator program has its own page, with some information that characterizes it, such as the services that provide and who are the beneficiaries of these services, as well the methodology, sector, among others. In addition to, the user can know more about the service through the website available on the page.

Courses	Link
Blockchain in Banking Industry & Enterprise Application	link
AI and Blockchain: A Disruptive Integration	link
Introduction to FinTech	link
Fintech: Blockchain for Business and Finance	link
Fintech: AI & Machine Learning in the Financial Industry	link
Deep Learning and Neural Networks for Financial Engineering	link
FinTech Ethics and Risks	link
Classical Machine Learning for Financial Engineering	link
Fintech: Overview of the Fintech Sector	link
Blockchain and FinTech: Basics, Applications, and Limitations	link
Introduction to Hyperledger Blockchain Technologies	link
Blockchain: Understanding Its Uses and Implications	link
Blockchain Technologies: Business Innovation and Application	link
UNIC's Free MOOC: Introduction to Decentralized Finance (DeFi)	link
MSc Financial Technology	link
Master in Financial Technology and Computing	link
EIT Digital Master School	link
Master in FinTech & Business Analytics EADA – ISDI	link
Data Strategy Executive Program	link
eBusines Executive Program	link
MsC in Digital Transformation	link
Fintech program	link
MsC in IoT & Data Science	link
IoT Executive Program	link
MsC in Data Science & Analytics	link
MsC in Business Analytics	link
MsC in Data Science & Deep Learning	link
IT/OT Security Awareness Training	link
High-quality training for computer security teams I	link
High-quality training for computer security teams II	link
BsC Business Information Technology	link
PhD Winter School	link
Embedded Security Engineering	link
Digitalization of Finance	link

Figure 2.10. Courses available on the INFINITECH marketplace.

Courses	Link
Certified Information Security Manager (CISM)	link
CRIPTOCERT Certification	link
MSc in Big Data	link
BLOCKCHAIN TECHNOLOGIES:BUSINESS INNOVATION AND APPLICATION	link
MSc in Computer Science, Concentration in Blockchain Technologies	link
MSc in Digital Currency	link
Retail and Digital Banking MSc	link
MSc in Fintech and Finance Innovation	link
MSc in Financial Technology (FinTech)	link

Figure 2.11. Courses available on the INFINITECH marketplace.

Workshops	Link
Decentralized Finance (DeFi) Webinar: Where do we stand and where do we go?	link
INFINITECH Stakeholders Workshop Series: "BigData and Artificial Intelligence for Portfolio Risk Assessment"	link
INFINITECH Stakeholders Workshops Series: "Artificial Intelligence and Big Data analytics applied to Personalised, Usage Based and Configurable Insurance Products"	link
INFINITECH Stakeholders Workshops Series: "Blockchain Applications for Digital Finance"	link
INFINITECH Stakeholders Workshops Series: "Risk Profiling and Portfolio Optimization for broader Use Cases"	link

Figure 2.12. Workshops on the INFINITECH marketplace.

Webinars	Link
Blockchain-enabled Consent Management	link
ERC1155 Token Smart Contract for Hyperledger	link
KYC/KYB On-Chain Data Governance	link

Webinars	Link
Tokenization on Hyperledger Fabric -ERC20 chaincode	link

Figure 2.13. Webinars on the INFINITECH marketplace.

To create community around the INFINITECH Marketplace and enrich it with new information, the following functionalities were added to allow the users interact with the platform:

- Social Login: The registration process was simplified and facilitated.
- Add new information: Any INFINITECH user can provide information.

Social Login

One of the objectives of the INFINITECH is to create a digital finance ecosystem of innovation, with IoT, Blockchain, BigData and AI solutions and services. For

this purpose, the Social Login was implemented in the INFINITECH Marketplace to connect the stakeholders and expand the INFINITECH community being an entry point for the community to access and provide information.

The Social login offers a simplified, quick, and easy registration, helping the users to sign-up on a third-party platform using their existing login information from social media networks, like Google, LinkedIn, and GitHub (Figure 2.15).

Add New Information

In order to enrich the digital finance ecosystem of innovation, it was created forms to give the users the opportunity to share their solutions and services on the INFINITECH Marketplace, because it is important that information continues to evolve and increase (See Figure 2.14).

Anyone can be part of this community, by registering, and as a INFINITECH user it is possible to provide information to the INFINITECH Marketplace. The users have different forms available, depending on the content they want to add (assets, courses, workshops, webinars, and accelerator programs).

Through the Homepage the users can access all the forms available in a single place (Figure 2.16) or go through the respective content page.

Accelerator Programmes	Link
Distributed Ledgers Research Centre (DLRC)	link
Gravity Ventures Incubator	link
Wayra	link
Visa Innovation Program	link
University of Valencia Science Park DIH	link
TVT Innovation – Pôle Mer Méditerranée	link
The LHoFT	link
Sunrise Valley Digital Innovation Hub (SV DIH)	link
StartupUtrecht	link
Startup Palace	link
Start-up Nation Central	link
Start-up garage	link
SmartCityTech	link
Seedrocket	link
Santaka Artificial Intelligence DIH	link
Portugal Fintech	link
Plug and Play	link
Nyuko a.s.b.l.	link
Novel-T	link

Figure 2.14. Continued

Accelerator Programmes	Link
Madeira Digital Innovation HUB	link
Luxinnovation	link
Le Pool French Tech Rennes St Malo	link
Lanzadera	link
IoT DIH	link
IRIS: European Digital Innovation Hub Navarra	link
AddedValue	link
Intelligent Urban Lab	link
INFOBALT DIH	link
Hub 612	link
Holland Fintech	link
Hellenic Blockchain Hub	link
Greek Fintech Hub	link
Frankfurt School Blockchain Center (FSBC) at the Frankfurt School of Finance & Management gGmbH (FS)	link
FinTech Scotland	link
FinTech Innovation Hub | Bank of Greece	link
Fintech District	link
Fintech Cluster	link
Fintech Aviv	link
Espaitec Science and Technology Park	link
Emprende UP	link
DigitalNorway	link
DIGIHALL	link
Cybersecurity Innovation HUB	link
CUATRECASAS ACELERA	link
Conector	link
Centre Val de Loire Digital Innovation Hub	link
Bridgeway Europe Startup Accelerator, Bridgeway Accelerator	link
BlackFin Tech	link
Bboster	link
AIR4S – Artificial Intelligence & Robotics for Sustainable Development Goals	link
AI4GALICIA: Artificial Intelligence for Galicia	link
Accelerator Frankfurt	link
Emerging Transactional and Financial Technology Hub (ETFTH)	link

Accelerator Programmes	Link
Insomnia	link
Copenhagen Fintech	link

Figure 2.14. Accelerator programs.

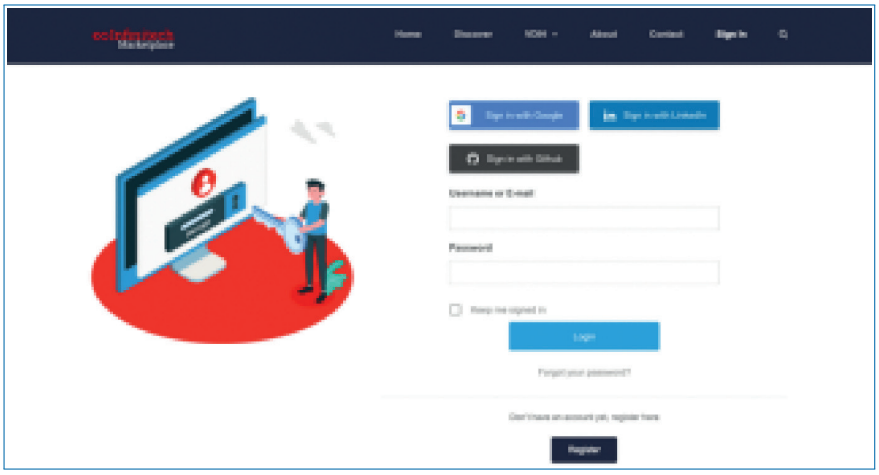


Figure 2.15. Social login.

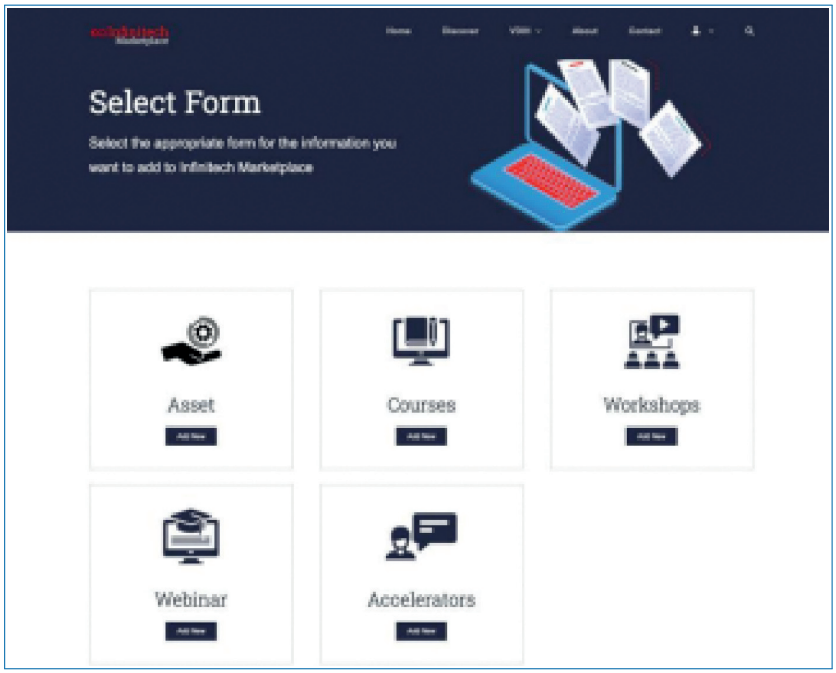


Figure 2.16. Select form.

Marketplace Usage Scenarios

The INFINITECH Marketplace is a good source of information, where can be found solutions and services about IoT, AI, Blockchain and BigData in context of finance/insurance. It is very important to enrich the information available and for

this the users will be capable of providing information, but also to take advantage of it.

So, this means, that can be seen from 2 perspectives, the user as an information provider and the user as information consumer. Both perspectives can be explored through the following scenarios that represent examples of the INFINITECH Marketplace usage:

- Scenario 1: A user who want add information to the INFINITECH Marketplace.
- Scenario 2: A user who want to know more about workshops in the finance domain.

Scenario 1: Upload a VDIH

Scenario 1 represents a user who want provide information to the INFINITECH Marketplace. For example, to upload a blockchain Course. For this, the user needs to follow a few steps:

- Step A: Sign in.
  - Step A1: If the user isn't registered, they will have to register first and then they will be able to login.
- Step B: Access to the form, which can be done through:
  - Step B1: Homepage, which gives access to all available forms.
  - Step B2: VDIH page, which is organized in "Training" and "Innovation Services". The user must go through Training to access the Courses page and select "Add New Course".
- Step C: Fill in the Course form. The first fields of the form refer to the provider (Your Information) and the other to the Course (Course Information). Some fields are mandatory, those that are identified with a*.
- Step D: When submitting the information, it is moderated by the consortium.
- Step E: The Course is published on the INFINITECH Marketplace.

Figure 2.17 show the aspect of a Course which was added to the INFINITECH Marketplace through Course form.

Scenario 2: Consult a VDIH

Scenario 2 represents a user that want to know more about blockchain Workshops in finance domain. The INFINITECH Marketplace has a VDIH section and in this section, the user can consult Workshops, but also other types of content are available about IoT, Blockchain, BigData, AI for finance and insurance (Courses, Webinars and Accelerator Programmed).

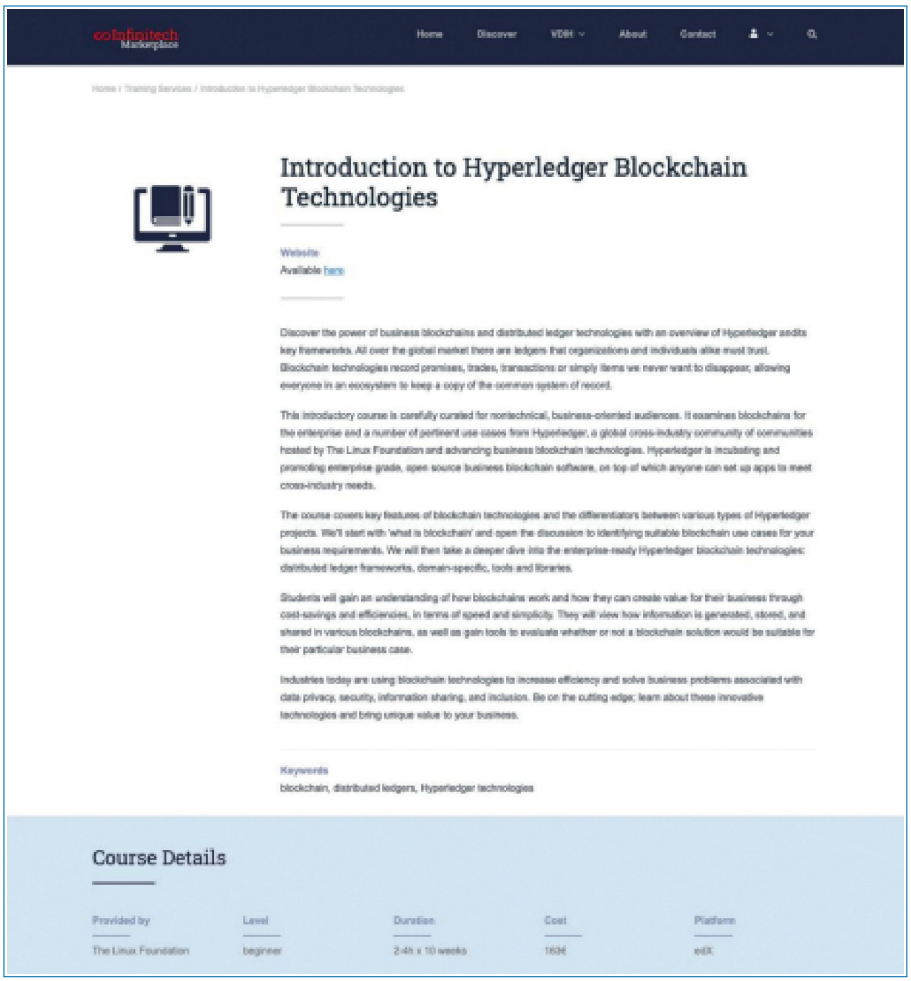


Figure 2.17. Introduction to hyperledger blockchain course.

On the Workshops page the user can get an idea of the information present on the platform through the names of the workshops and start by selecting one that seems to be of interest to him. On the respective workshop page, the user can find out more about the workshop, exploring the description and the agenda, among other additional information, and only follow the parts that fit their interests. The user can manage their time and later consult the presentation files that are available online.

To obtain more detailed information the user can contact the consortium and receive all the information that needs. Figure 2.18 represents an example of Workshops available at the INFINITECH Marketplace and we can see the sections it was mentioned before.

Figure 2.18. INFINITECH stakeholders workshops series: "Blockchain applications for digital finance".

2.3 Baseline Technologies

The back-end is the core base of the market platform and it has been developed using a variety of technologies/tools. First of all, its components are containerized in Docker images that, among others, offer more efficient management and maintenance, enabling continuous updates and integration. Python is used as the programming language that along with the Flask framework, which is a Web Server Gateway Interface (WSGI) developed in Python, implements RESTful APIs to handle the respective HTTP requests.

The offered assets are stored in a MongoDB No-SQL database that is used in combination with the file system of the hosting operating system for storing and retrieving large files/objects, of any format. The initial implementation contained the GridFS specification which was eventually deprecated as for really large files there was an extra delay on the back-end's responses for assets retrieval. Moreover, Gunicorn, a Python WSGI HTTP Server for UNIX, is utilized with NGINX, an open-source high-performance HTTP web server and reverse proxy, since Flask is not optimum for production mode, and thus, both tools will extend the Flask framework in order to enable access to multiple users at the same time.

In terms of the front-end, it has been implemented using various web technologies (HTML, CSS, etc.) and it is functional using PHP and JavaScript technologies. It also exploits WordPress and various plugins of it, in order to manage the content that is presented. [D8.4]

2.4 Interfaces (APIs)

This section describes the REST API endpoints that are introduced in the final version of the back-end. As already mentioned, the back-end is a REST API that receives HTTP requests to trigger its designed and implemented functionalities. These APIs are categorized into 3 main groups, namely: (1) APIs related to Users, (2) APIs related to Descriptions and (3) APIs related to Assets.

2.4.1 APIs Related to Users

This group of APIs offers functionalities intended for the management of marketplace users. The most important functionality is that of user registration, as it is necessary for the usage of the rest functionalities. For all users, except for their personal information, there will be a unique username. The Table 2.1 presents the endpoints related to Users, as they are in the first version of the marketplace.

Table 2.1. APIs of the back-end related to Users.

Action	HTTP Method	Endpoint
Register a new user (Sign up)	POST	{HOST}/accounts/users/registration
Check the availability of a username	GET	{HOST}/accounts/username/availability
Authenticate a user (Login)	POST	{HOST}/accounts/users/authentication
Get user's information	GET	{HOST}/accounts/users/information/{username}
Update user's information	PUT	{HOST}/accounts/users/information/{username}
Change user's password	POST	{HOST}/accounts/users/password/change
Reset user's password	POST	{HOST}/accounts/users/password/reset
Delete user's account	DELETE	{HOST}/accounts/users/delete/{username}
Get user's privacy data	GET	{HOST}/accounts/users/data
Erase user's data	DELETE	{HOST}/accounts/users/data
Set user's profile as public/private	PUT	{HOST}/accounts/users/profile/accessibility
Connect user's account to a social service	POST	{HOST}/accounts/users/sso/{sso_service}/connect
Disconnect user's account from a social service	POST	{HOST}/accounts/users/sso/{sso_service}/disconnect

- *{HOST}* refers to the hosting server: the domain name and the port that the back-end runs.
- *{username} refers to "user's unique username".*
- {*sso_service*} refers to a supported SSO provider (i.e. Google, LinkedIn or GitHub).
- Some of these actions require additional fields in the headers or even in the body of the HTTP request. Example of a required field is the API key that users use in order to validate themselves to the platform.

A more detailed description of all the Interfaces listed in the previous table is presented at the Table 2.2 to 2.14, a table per Interface on the list respectively.

Table 2.2. Interface: Register a new user (Sign up).

Title:	Register a new user (Sign up)
Endpoint:	{HOST}/accounts/users/registration
HTTP Method:	POST
Description:	From this endpoint, the registrations of the marketplace users are made. A POST request should be submitted and the next JSON schema must be in its body as raw data. It is noted that a) the email and the username must be unique and available b) the schema below should be exactly the same, whether there are values or not (empty strings "") – the array "social" can be empty.

(*Continued*)

Table 2.2. Continued

	`{ "username": "…", "account": {"password": "…"}, "info": { "first_name": "…", "last_name": "…", "email": "…", "about": "…" } }`	
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	None	
Restrictions/Special Features:	None	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request POST '{HOST}/accounts/users/registration' \
--header 'Content-Type: application/json'
--data-raw '{
  "username": "…", "account": {"password": "…"},
  "info": {
      "first_name": "…", "last_name": "…",
      "email": "…", "about": "…"
  }
}'
```

After a successful registration, the following JSON document is stored in the database:

```
{
  "username": "…", // user's unique username
  "account": {
      "password": "…", // user's password (hashed)
      "password_protected": "…", // parameter that determines
      whether the account is      password protected or not
      (values 1 or 0)
      "connections": {"google": "…", … } // object that
      determines if the account is  connected to any of the
      supported SSO services (e.g. Google, Github, LinkedIn,
      etc.)
      "role": "user", // user's role
      (user or admin) "verified": "…",
      // value = "1" if user is
      verified,
      otherwise, it has a verification code to use it for user's
      email/account verification
      "registration_datetime": "…" // user's registration date
      "public_profile": "…" // parameter that determines whether
      the account can be      publicly displayed or not (values
      1 or 0)
  },
  "info": {// info provided during user's registration
      "first_name": "…", "last_name": "…", "email": "…",
      "about": "…"
  }
}
```

Table 2.3. Interface: Check the availability of a username.

Title:	Check the availability of a username	
Endpoint:	{HOST}/accounts/username/availability	
HTTP Method:	GET	
Description:	This endpoint is used in order to check the availability of a username during the users' registration. A GET request should be made and the key "x-username" must be included in the headers of the request.	
Body Data:	None	
Headers:	Key	Value
	x-username	The username whose availability will be checked.
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	None	
Restrictions/Special Features:	None	
Successful Response:	Availability status in JSON Object.	

The following is an example of the request in cURL:

```
curl --request GET '{HOST}/accounts/username/availability' -header 'x-
username: <value>'
```

Table 2.4. Interface: Authenticate a user (Login).

Title:	Authenticate a user (Login)	
Endpoint:	{HOST}/accounts/users/authentication	
HTTP Method:	POST	

Description:	Through this endpoint, the users are authenticated in order to log in to their account. A POST request should be made and the next JSON schema, containing users' credentials, must be in the body of the request as raw data. It is noted that users can log in either with their email or with their username. Finally, single sign-on (SSO) schemes are supported (only through the front-end) for registration and login, using accounts from social media (i.e. Google, LinkedIn, Github).
	`{ "username": "…", "email": "…", "password": "…" }'`
	A successful response will return the next JSON schema that contains the API key in the key "token":
	`{"_status": "successful", "token": "<api_key>"}`	

Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	None	
Restrictions/Special Features:	None	
Successful Response:	JSON Object with a successful message and the API key.	

The following is an example of the request in cURL:

```
curl --request POST
'{HOST}/accounts/users/authentication' \ -
-header 'Content-Type: application/json' \
--data-raw '{ "username": "…", "email": "…", "password": "…" }'
```

Table 2.5. Interface: Get user's information.

Title:	Get user's information
Endpoint:	{HOST}/accounts/users/information/{username}
HTTP Method:	GET
Description:	This endpoint is used in order to retrieve information about a user. Some of the information that is retrieved, are first and last name, about section, email, etc (below is illustrated an example of retrieved user's information). A GET request should be made and the user's {username} is required at the end of the endpoint. Moreover, this endpoint is restricted and thus, requesters' API keys must be included in the headers of the request.
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)
Headers:	Key: Value API_KEY: Requester's API key.
URL Parameters:	Parameter: Value username: The username of the user whose information will be retrieved.
Query Parameters:	None
Restrictions/Special Features:	The administrators and the accounts' owners are able to retrieve all users' information, while users that retrieve information of other users retrieve only public information.
Successful Response:	JSON Object with a user's information.

The following is an example of the request in cURL:

```
curl --request GET '{HOST}/accounts/users/information/{username}' --
header 'API_KEY: <value>'
```

Example of a successful response

```
{"_status": "successful", "result": {
        "account": {"registration_datetime": "…", "role": "user",
        "verified": "1",
                    "connections": {"google": "…", … }
        },
        "info": {"about": "…", "email": "…", "first_name": "…", "last_name": "…"},
        "username":
"…" }}
```

Table 2.6. Interface: Update user's information

Title:	Update user's information
Endpoint:	{HOST}/accounts/users/information/{username}
HTTP Method:	PUT
Description:	This endpoint handles requests for updating users' information. A PUT request should be made and the next JSON schema (it is flexible and thus may contain fewer fields – but without new fields), containing users' new information, must be in its body as raw data. `{ "first_name": "…", "last_name": "…", "about": "…", "email": "…", …  }`

(*Continued*)

Table 2.6. Continued

	Moreover, this endpoint is restricted and thus, users' API keys must be included in the headers of the request. It should be noted that only the accounts' owners and the administrators are able to update the information of a user.
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)
Headers:	Key: API_KEY; Value: Requester's API key.
URL Parameters:	Parameter: username; Value: The username of the user whose information will be updated.
Query Parameters:	None
Restrictions/Special Features:	Only the accounts' owners and the administrators are able to update the information of a user.
Successful Response:	A successful response will return the next JSON Object that contains a new API key in the key "token": `{"_status": "successful", "message": "The information of the user with username '{username}' has been updated.", "token": "<api_key>"}`

The following is an example of the request in cURL:

```
curl --request PUT
'{HOST}/accounts/users/information/{username}' \ -
-header 'API_KEY: <value>' --header 'Content-Type:
application/json' \
--data-raw '{ "first_name": "…", "last_name": "…", "about": "…",
"email": "…" }'
```

Table 2.7. Interface: Change user's password.

Title:	Change user's password
Endpoint:	{HOST}/accounts/users/password/change
HTTP Method:	POST
Description:	This endpoint is used when the users want to change their accounts' password. A POST request should be made and the next JSON schema, containing users' new and old credentials, must be in its body as raw data. Also, this endpoint is restricted and thus, users' API keys must be included in the headers of the request. It should be noted that this action is only available to accounts' owners. `{"old_password": "…", "new_password": "…", "confirm_new_password": "…"}`
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)
Headers:	Key: API_KEY; Value: Requester's API key.
URL Parameters:	Parameter: None; Value: None
Query Parameters:	None

(*Continued*)

Table 2.7. Continued

Restrictions/Special Features:	Only available to accounts' owners. The new password must not be the same with previous password.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request POST '{HOST}/accounts/users/password/change' \
--header 'API_KEY: <value>' --header 'Content-Type: application/json' \
--data-raw '{ "old_password": "…", "new_password": "…",
"confirm_new_password": "”}'
```

Table 2.8. Interface: Reset user's password.

Title:	Reset user's password
Endpoint:	{HOST}/accounts/users/password/reset
HTTP Method:	POST
Description:	This endpoint handles the process of changing users' passwords after a password reset request. It works in combination with the front-end's features which, at first, sends an email to the users with a password reset link that redirects to a form from which the users can set their new password. After this process, the front-end sends a POST request to the back-end with the new credentials of the user (username and password). `{ "username ": "…", "password": "…", "password_reset_code": "…" }`
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)
Headers:	Key / Value None / None
URL Parameters:	Parameter / Value None / None
Query Parameters:	None
Restrictions/Special Features:	None
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl -request POST '{HOST}/accounts/users/password/reset' \
--header 'Content-Type: application/json' \
--data-raw '{ "username ": "…", "password": "…", "password_reset_code":
"…" }'
```

Table 2.9. Interface: Delete user's account.

Title:	Delete user's account
Endpoint:	{HOST}/accounts/users/delete/{username}
HTTP Method:	DELETE
Description:	In order to delete an account, this endpoint should be used, making a DELETE request and providing requester's password in its body, as raw data (JSON format). The endpoint must contain the username of the user whose account will be deleted at the end of the URL. `{ "password": "…" }`

(*Continued*)

Table 2.9. Continued

	The endpoint is restricted and thus, requester's API key must be included in the headers of the request. This action is available to accounts' owners and to administrators who are able to delete users from the marketplace. If the action is made by an administrator, the value of the field "password" in the body should be the password of administrator. An important note is that the deletion of an account has as result the deletion of all user's data, offered descriptions and assets.
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)
Headers:	Key: API_KEY Value: Requester's API key.
URL Parameters:	Parameter: username Value: The username of the user whose account will be deleted.
Query Parameters:	None
Restrictions/Special Features:	Only the accounts' owners and the administrators are able to delete an account/user.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request DELETE
'{HOST}/accounts/users/delete/{username}' \ --
header 'API_KEY: <value>' --header 'Content-
Type: application/json' \
--data-raw '{ "password": "…" }'
```

Table 2.10. Interface: Get user's privacy data.

Title:	Get user's privacy data
Endpoint:	{HOST}/accounts/users/data
HTTP Method:	GET
Description:	This endpoint, which is available only to accounts' owners, returns all the personalized data of the requester, returning users' information and various metadata. The endpoint is restricted and thus, the API key of the requester must be included in the headers of the request.
Body Data:	None
Headers:	Key: API_KEY Value: Requester's API key.
URL Parameters:	Parameter: None Value: None
Query Parameters:	None
Restrictions/Special Features:	Available only to accounts' owners.
Successful Response:	A JSON Object with users' data.

The following is an example of the request in cURL:

```
curl --request GET '{HOST}/accounts/users/data' --header 'API_KEY:
<value>'
```

Table 2.11. Interface: Erase user's data.

Title:	Erase user's data
Endpoint:	{HOST}/accounts/users/data
HTTP Method:	DELETE
Description:	This endpoint, which is available only to accounts' owners, erases all the personalized data of the requester. The endpoint is restricted and thus, the API key of the requester must be included in the headers of the request.
Body Data:	None
Headers:	Key: API_KEY; Value: Requester's API key.
URL Parameters:	Parameter: None; Value: None
Query Parameters:	None
Restrictions/Special Features:	Available only to accounts' owners.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/accounts/users/data' --header 'API_KEY:
<value>'
```

Table 2.12. Interface: Set user's profile as public/private.

Title:	Set user's profile as public/private
Endpoint:	{HOST}/accounts/users/profile/accessibility
HTTP Method:	PUT
Description:	This endpoint, which is available only to accounts' owners, turns the user's account as publicly accessible or private, depending on the preferences of the user. The indication of the account's accessibility is included in the headers of the request under the "accessibility" key. Moreover, the endpoint is restricted and thus, the API key of the requester must be included in the headers of the request.
Body Data:	None
Headers:	Key: API_KEY; Value: Requester's API key. Key: accessibility; Value: Determines whether the account will be accessible or not. Acceptable values: • "1": public account • "0": private account
URL Parameters:	Parameter: None; Value: None
Query Parameters:	None
Restrictions/Special Features:	Available only to accounts' owners.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/accounts/users/profile/accessibility' \
--header 'API_KEY: <value>' --header 'accessibility: <value>'
```

Table 2.13. Interface: Connect user's account to a social service.

Title:	Connect user's account to a social service
Endpoint:	{HOST}/accounts/users/sso/{sso_service}/connect
HTTP Method:	POST
Description:	This endpoint, which is available only to accounts' owners, connects an existing marketplace account to a social service through the single sign-on (SSO) functionality. This action is not a sign up/ registration action and thus, the action of the registration can be done only through the "User's registarion" interface (Table 2.2). The current supported SSO services are the following: Google, LinkedIn and GitHub. The body of the request must contain the unique ID of the user in the corresponding {sso_service} with which the connection will be made. `{ "id": "…" }` Moreover, the endpoint is restricted and thus, the API key of the requester must be included in the headers of the request.
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json)
Headers:	Key: Value API_KEY: Requester's API key.
URL Parameters:	Parameter: Value sso_service: The name of the SSO service. Acceptable values: "google", "linkedin", "github".
Query Parameters:	None
Restrictions/Special Features:	Available only to accounts' owners.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request POST
'{HOST}/accounts/users/sso/{sso_service}/connect' \ --
header 'API_KEY: <value>' --header 'Content-Type:
application/json' \
--data-raw '{ "id": "…" }'
```

Table 2.14. Interface: Disconnect user's account from a social service.

Title:	Disconnect user's account from a social service
Endpoint:	{HOST}/accounts/users/sso/{sso_service}/disconnect
HTTP Method:	POST
Description:	This endpoint, which is available only to accounts' owners, disconnects a social account/SSO service from the requester's account. This action is possible only if the users have set a password for their accounts. If their accounts are not password protected, then the users must set a password through the "reset password" interface (Table 2.8), before disconnecting the services from their accounts. The current supported SSO services are the following: Google, LinkedIn and GitHub. Moreover, the endpoint is restricted and thus, the API key of the requester must be included in the headers of the request.
Body Data:	None
Headers:	Key: Value API_KEY: Requester's API key.

(*Continued*)

Table 2.14. Continued

URL Parameters:	Parameter	Value
	sso_service	The name of the SSO service. Acceptable values: "google", "linkedin", "github".
Query Parameters:	None	
Restrictions/Special Features:	Available only to accounts' owners.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request POST
'{HOST}/accounts/users/sso/{sso_service}/disconnect' \ --
header 'API_KEY: <value>'
```

2.4.2 APIs Related to Descriptions

This group of APIs offers functionalities intended for the management of the descriptions. They support all CRUD operations as well as the search functionality. Special emphasis was placed on the APIs for the descriptions' retrieval, extending them so as to get the latest descriptions or even random descriptions either from a specific collection (database collection) or from all collections at once, using a keyword named "all". The collections of the database, as described in deliberable D8.2 "Market Platform and VDIH Specifications – II", vary as well as the offered marketplace's types of assets. The current list of the collections can be found at the end of the following Table 2.15, which presents the endpoints related to Descriptions. The details of each of Interfaces can be found at Table 2.16 to 2.37 respectively.

Table 2.15. APIs of the back-end related to Descriptions.

Action	HTTP Method	Endpoint
Get a list with all descriptions	GET	{HOST}/descriptions/all
Get a list with all descriptions from a specific collection	GET	{HOST}/descriptions/{collection}
Get a specific description (using keyword "all")	GET	{HOST}/descriptions/all/{description_id}
Get a specific description (using description's "collection")	GET	{HOST}/descriptions/{collection}/{description_id}
Get the latest descriptions from all collections	GET	{HOST}/descriptions/all/latest
Get the latest descriptions from a specific collection	GET	{HOST}/descriptions/{collection}/latest
Get random descriptions from all collections	GET	{HOST}/descriptions/all/random

(*Continued*)

Table 2.15. Continued

Action	HTTP Method	Endpoint
Get random descriptions from a specific collection	GET	{HOST}/descriptions/{collection}/random
Upload/Create a new description with random ID	POST	{HOST}/descriptions/{collection}
Upload/Create a new description with given ID	POST	{HOST}/descriptions/{collection}/{given_id}
Update a specific description (using keyword "all")	PUT	{HOST}/descriptions/all/{description_id}
Update a specific description (using description's "collection")	PUT	{HOST}/descriptions/{collection}/{description_id}
Delete a specific description (using keyword "all")	DELETE	{HOST}/descriptions/all/{description_id}
Delete a specific description (using description's "collection")	DELETE	{HOST}/descriptions/{collection}/{description_id}
Delete all descriptions	DELETE	{HOST}/descriptions/all/all
Delete all descriptions from a specific collection	DELETE	{HOST}/descriptions/{collection}/all
Get a list with all descriptions that need permission (administrators' action)	GET	{HOST}/descriptions/permit/all
Get a list with all descriptions from a specific collection that need permission (administrators' action)	GET	{HOST}/descriptions/permit/{collection}
Approve or reject a description that needs permission, using keyword "all" (administrators' action)	POST	{HOST}/descriptions/permit/all/{description_id}
Approve or reject a description that needs permission, using description's "collection" (administrators' action)	POST	{HOST}/descriptions/permit/{collection}/ {description_id}
Approve or reject all descriptions that need permission, using keyword "all" (administrators' action)	POST	{HOST}/descriptions/permit/all/all
Approve or reject all descriptions that need permission under a specific collection, using a "collection" value (administrators' action)	POST	{HOST}/descriptions/permit/{collection}/all

- *{HOST}* refers to the hosting server: the domain name and the port that the back-end runs.
- *{description_id}* refers to the ID of a specific description.
- *{given_id}* is used in "upload description" action, providing new-description's ID.
- As a {collection} can be one of the following values derived from the current types of offered assets:
  {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications",
  "blockchain", "third-party-tools", "whitepapers", "how-to-videos", "accelerator-programmes",
  "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
- Some of these actions require additional fields in the headers or even in the body of the HTTP request. Example of a required field is the API key that users use in order to validate themselves to the platform.

Table 2.16. Interface: Get a list with all descriptions.

Title:	Get a list with all descriptions
Endpoint:	{HOST}/descriptions/all
HTTP Method:	GET
Description:	A GET request to this endpoint will result in the retrieval of the stored descriptions from all collections. It uses the keyword "all" instead of a specific collection, which makes the platform to retrieve descriptions from all collections at once. The descriptions that return from this request are in a short schema (short description), meaning that the retrieved information is limited. An example of a description's description in short schema is the following JSON schema:

```
{"collection": "datasets", "id": "datasets_vlLZWaoQN1Fe ",
  "info": { "title": "Example title.",
      "keywords": [ "information"], "owner": "Vasilis Koukos
      "short_desc": "This is an example", "type": "dataset"
  }, metadata": {"provider": "vkoukos",
       "updateDate": "…", "uploadDate": "…", "views": 35
}
}
```

This endpoint can get query parameters to search for descriptions that meet certain conditions. As a query parameter can be any pair of key-value, while additional search operators can be used for more advanced and enhanced search. More details about searching can be found in Section 3.2.3. Also, this endpoint offers some standard query parameters, as described below (Query Parameters).

Body Data:	None
Headers:	Key Value
	None None

(*Continued*)

Table 2.16. Continued

URL Parameters:	Parameter	Value
	None	None
Query Parameters:	Key	Value
	sortBy	[Optional] Sorts the descriptions by a field – the default is the "newest" key. The value should be one of the following: "**newest**": sort by date in descending order. "**oldest**": sort by date in ascending order. "**views-asc**": sort by the number of views in ascending order. "**views-desc**": sort by the number of views in descending order. "**title**": sort by title in ascending order.
	itemsPerPage	[Optional] Returns the results separated in pages (arrays) of N items. The number N is specified by the value of this key. The value N must be an integer number greater or equal to 1. If the key is not used or has a non-accepted value, the results are returned on a single page.
	page	[Optional] This key can only be used if the "itemsPerPage" key is also used. If it is used, it returns only the specified (by key's value) page instead of all pages created using the key "itemsPerPage". The value must be an integer number greater or equal to 1. The default value is 0, which means that all pages will be returned.
	Any key to search (refer to Section 3.2.3)	Any value to search (refer to Section 3.2.3).
Restrictions/Special Features:	None	
Successful Response:	A JSON Object with the results (all descriptions from all collections). If the query parameter "itemsPerPage" is used, then the results contain the total number of the pages.	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/all'
+ curl --request GET '{HOST}/descriptions/all?sortBy={value}'
+ curl --request GET '{HOST}/descriptions/all?itemsPerPage={value}'
+ curl --request GET
'{HOST}/descriptions/all?itemsPerPage={value}&page={value}'
+ curl --request GET
'{HOST}/descriptions/all?sortBy={value}&itemsPerPage={value}'
+ curl --request GET
'{HOST}/descriptions/all?sortBy={value}&itemsPerPage={value}&pa
ge={value}'
Example of retrieving the 10 most viewed descriptions:
+ curl --request GET '{HOST}/descriptions/all?sortBy=views-
desc&itemsPerPage=10&page=1'
```

Table 2.17. Interface: Get a list with all descriptions from a specific collection.

Title:	Get a list with all descriptions from a specific collection	
Endpoint:	{HOST}/descriptions/{collection}	
HTTP Method:	GET	
Description:	This GET request is similar to the above request. The only difference between these two actions is that this request retrieves descriptions from a single – specific collection (instead of using keyword "all"). For more details, refer to the previous endpoint.	
Body Data:	None	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
Query Parameters:	As in the above request.	
Restrictions/Special Features:	None	
Successful Response:	A JSON Object with the results (all descriptions in a specific collection).	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/{collection}'
+ curl --request GET \

'{HOST}/descriptions/{collection}?sortBy={value}&itemsPerPage={value}&p
age={value}'
```

Table 2.18. Interface: Get a specific description (using keyword "all").

Title:	Get a specific description (using keyword "all")
Endpoint:	{HOST}/descriptions/all/{description_id}
HTTP Method:	GET
Description:	With this GET request, the users are able to retrieve a specific description. The retrieval of a specific description is possible using its unique identification code (ID) that is known when uploading it. Also, the retrieval of a specific description can be done using both keyword "all" and the name of the collection that the description has been stored (next interface). This is feasible because the back-end ensures that the IDs are unique regardless of the collection a description has been stored. Moreover, the retrieval of a specific description requires an API key to be retrieved in its "full schema". If requester's API key is missing, then the endpoint returns the short schema of the description's description. Example of a full schema of a description is in the endpoint that handles the uploading of a description (Table 2.24).
Body Data:	None

(*Continued*)

Table 2.18. Continued

Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	description_id	The ID of the description that will be retrieved.
Query Parameters:	None	
Restrictions/Special Features:	The full schema is available only to authenticated (and verified) users, otherwise, the short schema is available to all.	
Successful Response:	A JSON Object with the description in the results.	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/all/{description_id}'
+ curl --request GET '{HOST}/descriptions/all/{description_id}' --
header 'API_KEY: <value>'
```

Table 2.19. Interface: Get a specific description (using description's "collection").

Title:	Get a specific description (using description's "collection")	
Endpoint:	{HOST}/descriptions/{collection}/{description_id}	
HTTP Method:	GET	
Description:	This GET request is similar to the above request, with the difference that it uses description's collection for the retrieval of a description (instead of using keyword "all"). The value of the {collection} must be the collection in which the description has been stored. More information about the endpoint can be found in the previous endpoint.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	description_id	The ID of the description that will be retrieved.
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "third-party-tools", "whitepapers", "how-to-videos", "accelerator-programmes", "innovation-support-services", "workshops", "courses", "containers", "uncategorized", "other"}
Query Parameters:	None	
Restrictions/Special Features:	The full schema is available only to authenticated (and verified) users, otherwise, the short schema is available to all.	
Successful Response:	A JSON Object with the description in the results.	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/{collection}/{description_id}'
+ curl --request GET '{HOST}/descriptions/{collection}/{description_id}'
\
        --header 'API_KEY: <value>'
```

Table 2.20. Interface: Get the latest descriptions from all collections.

Title:	Get the latest descriptions from all collections	
Endpoint:	{HOST}/descriptions/all/latest	
HTTP Method:	GET	
Description:	This request is used to retrieve the most recent uploaded descriptions sorted based on the date that they have been uploaded, with the most recent being on the top of the list. This is a GET request, using the keyword "all" and returns the K latest descriptions from all collections. The value of K can be specified through the query parameter "max" (the default value is 20). The descriptions are returned in their short schema. Finally, the endpoint "Get a list with all descriptions" can return the same results as the current, if the example at the end of the current endpoint will be followed.	
Body Data:	None	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	Key	Value
	max	Integer value greater than 0 – Default: 20
Restrictions/Special Features:	None	
Successful Response:	A JSON Object with the results (latest descriptions from all collections).	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/all/latest'
+ curl --request GET '{HOST}/descriptions/all/latest?max=5'


Example of similar response by the endpoint "Get a list with all
descriptions":
+ curl --request GET
'{HOST}/descriptions/all?sortBy=newest&itemsPerPage=20&page=1'
```

Table 2.21. Interface: Get the latest descriptions from a specific collection.

Title:	Get the latest descriptions from a specific collection
Endpoint:	{HOST}/descriptions/{collection}/latest
HTTP Method:	GET
Description:	This request is similar to the above GET request. It uses the value of a specific collection and not the keyword "all", which results to return sorted the K most recent descriptions of the provided collection. The value of K can be specified through the query parameter "max" (the default value is 20). The descriptions are returned in their short schema. Finally, the endpoint "Get a list with all descriptions from a specific collection" can return the same results as the current, if the example at the end of the current endpoint will be followed.
Body Data:	None

(*Continued*)

Table 2.21. Continued

Headers:	Key	Key
	None	None
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "third-party-tools", "whitepapers", "how-to-videos", "accelerator-programmes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
Query Parameters:	Key	Value
	max	Integer value greater than 0 – Default: 20
Restrictions/Special Features:	None	
Successful Response:	A JSON Object with the results (latest descriptions of a collection).	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/{collection}/latest'
+ curl --request GET '{HOST}/descriptions/{collection}/latest?max=5'


Example of similar response by the endpoint "Get a list with all

collection":
descriptions from a specific
+ curl --request GET
'{HOST}/descriptions/{collection}?sortBy=newest&itemsPerPage=20&page
=1'
```

Table 2.22. Interface: Get random descriptions from all collections.

Title:	Get random descriptions from all collections	
Endpoint:	{HOST}/descriptions/all/random	
HTTP Method:	GET	
Description:	This endpoint returns a number of random descriptions from all collections (uses keyword "all"). It is useful in order to suggest and promote different descriptions each time. It is also used in the home page of the INFINITECH market platform, where random descriptions are displayed. Through the query parameter "max" it can return K descriptions, where K can be specified by the users (the default value is 4). The descriptions are returned in their short schema.	
Body Data:	None	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	Key	Value
	max	Integer value greater than 0 – Default: 20
Restrictions/Special Features:	None	

(*Continued*)

Table 2.22. Continued

Successful Response:	A JSON Object with the results (random descriptions from all collections).

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/all/random'
+"curl --request GET '{HOST}/descriptions/all/random?max=5'
```

Table 2.23. Interface: Get random descriptions from a specific collection.

Title:	Get random descriptions from a specific collection	
Endpoint:	{HOST}/descriptions/{collection}/random	
HTTP Method:	GET	
Description:	This endpoint is similar to the above endpoint. Instead of keyword "all" it uses a specific collection and thus it returns a number of K random descriptions of the provided specific collection. The value of K can be specified through the query parameter "max" (the default value is 4). The descriptions are returned in their short schema.	
Body Data:	None	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "third-party-tools", "whitepapers", "how-to-videos", "accelerator-programmes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
Query Parameters:	Key	Value
	max	Integer value greater than 0 – Default: 20
Restrictions/Special Features:	None	
Successful Response:	A JSON Object with the results (random descriptions from a specific collection).	

The following is an example of the request in cURL:

```
+ curl --request GET '{HOST}/descriptions/{collection}/random'
+ curl --request GET '{HOST}/descriptions/{collection}/random?max=5'
```

Table 2.24. Interface: Upload/Create a new description with random ID.

Title:	Upload/Create a new description with random ID
Endpoint:	{HOST}/descriptions/{collection}
HTTP Method:	POST

(*Continued*)

Table 2.24. Continued

Description: Through this POST request, the users can upload their descriptions. It requires users/providers to specify (at the end of the endpoint) the collection in which the description will be stored. Also, in the headers of the request, the API key of the provider is necessary to be included because the endpoint is available only to authenticated (and verified) users. An important note is that all the new descriptions uploaded to the Marketplace must be approved by an administrator before they can be made available to other users. Moreover, the administrators can upload a description on behalf of other users, adding the key "x-provider" in the headers of the request. The body of the request must contain the contents of the description as raw data in JSON format. The schema of the descriptions' content varies, and it is flexible to be extended. The JSON schema below, presents some required fields of a description:

```
{
   "title": "<title of the asset>",
   "description ": "<description of the
provided asset>",
   "type": "<type of the asset (same as the
collection value)>",
   "owner ": "<organization / author /
etc.>",
   "contact": "<provider's name and email>",
   "availability": "<reflects the users who
can see the description:          public /
infinitech / specific Work Packages or
Pilots / Other>",    "keywords": ["<keyword
1>", … ], "comments": "<provider's
comments>" }
```

Except of these fields, there are also some optional fields, like "subtype", "deliveryDate", "resources", "fieldOfUse", "license" and others. The "resources" field is used in cases that the descriptions contain assets that are stored in other repositories (e.g. GitHub, Gitlab, etc.).
The front-end has appropriate forms that build automatically these JSON schemas.

Body Data: Raw (JSON) Data – as the above schema (Content-Type: application/json).
It should be noted that the descriptions can also be uploaded from binary files that contain the above JSON schema (a curl example can be found below).

Headers:

Key	Value
API_KEY	Requester's API key.
x-provider	[Optional & only for administrators] The username of the provider user in case that the description is uploaded by an administrator and not by the provider.

URL Parameters:

Parameter	Value
collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses", "containers", "uncategorized", "other"}

(*Continued*)

Table 2.24. Continued

Query Parameters:	None
Restrictions/Special Features:	Available to all authenticated (and verified) users. The administrators can upload a description on behalf of other users.
Successful Response:	JSON Object with the new description's ID in its content.

The following is an example of the request in cURL:

```
curl -request POST '{HOST}/descriptions/{collection}' \
--header 'API_KEY: <value>' -header 'Content-Type:
application/json' \ --data-raw '{
   "title": "<title of the asset>",
   "description": "<description of the provided asset>",
   "type": "<type of the asset (same as the collection value)>",
   "owner ": "<organization / author / etc.>",
   "contact": "<provider's name and email>",
   "availability": "<reflects the users who can
see the description:          public / infinitech
/ specific Work Packages or Pilots / Other>",
"keywords": ["<keyword 1>", … ], "comments":
"<provider's comments>"
}
'
Example of uploading a description through binary
data/file: curl -request POST
'{HOST}/descriptions/{collection}' --header 'API_KEY:
<value>' \
--header 'Content-Type: application/json' --data-binary
'@<path_to_json_file>'
```

Below are some examples of the stored descriptions' schema:

Example 1 – Newly uploaded and updated description with no assets or resources

```
{
   "id": "algorithms_P8fYOAX67HkK-8fpe1TlB-KuR4-Zsck",
   "collection": "datasets",
   "info": {
            "title": "Example.", "type": "algorithms",
            "comments": "Private comment.",
            "contact": "Vasilis Koukos, email",
            "description": "This is an example of description.",
            "keywords": [ "testing", "documentation"]"],
            "owner": "UPRC", "availability": "public",
            // optional fields
            "subtype": "ML algorithm", "fieldsOfUse": ["machine learning",
"big data"],
            "license": "-", "any_other_field_provided_by_the_user": "...",
...
   },
   "metadata": {"approved": 1, //0 for pending / 1 for approved
                "last_updated_by": "vkoukos", "md5": "<md5 hash of the
description's data>",                  "provider": "vkoukos",
                "updateDate": "2022-11-15 13:50:48.420Z", "uploadDate":
"2022-10-13 10:34:28.420Z",
                "version": 2, //the version of the description -
increases when updating                  "views": 34},
    "assets": [],   //list with the uploaded assets for this description
    "resources": []  //list with the external resources / links
added to this description }
```

Example 2 – Description with an uploaded asset

```
{
   "id": "algorithms_P8fYOAX67HkK-8fpe1TlB-KuR4-Zsck",
   … … …
```

Table 2.25. Continued

Description:	This endpoint is similar to the above POST request. The only difference is that through the current endpoint, the users are able to specify the ID of the new description, providing it at the end of the endpoint {given_id}. Currently, this endpoint can be used only by the administrators.	
Body Data:	Raw (JSON) Data – as the schema of the previous endpoint (Content-Type: application/json). It should be noted that the descriptions can also be uploaded from binary files that contain the JSON schema of the previous endpoint (example in curl can be found at the end of the interface).	
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-provider	[Optional & only for administrators] The username of the provider user in case that the description is uploaded by an administrator and not by the provider.
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
	given_id	The ID to be given to the new description.
Query Parameters:	None	
Restrictions/Special Features:	Available only to administrators. The administrators are able to upload a description on behalf of other users.	
Successful Response:	JSON Object with the new description's ID in its content.	

The following is an example of the request in cURL:

```
curl -request POST
'{HOST}/descriptions/{collection}/{given_id}'
\ --header 'API_KEY: <value>' -header
'Content-Type: application/json' \ --data-raw
'{
   "title": "<title of the asset>",
   "description": "<description of the provided asset>",
   "type": "<type of the asset (same as the collection value)>",
   "owner ": "<organization / author / etc.>",
   "contact": "<provider's name and email>",
   "availability": "<reflects the users who can
see the description:          public /
infinitech / specific Work Packages or Pilots /
Other>",    "keywords": ["<keyword 1>", … ],
"comments": "<provider's comments>"
}
'
Example of uploading a description through binary data/file:
curl -request POST '{HOST}/descriptions/{collection}/{given_id}'
--header 'API_KEY: <value>' \
--header 'Content-Type: application/json' --data-binary
'@<path_to_json_file>'
```

Table 2.26. Interface: Update a specific description (using keyword "all").

Title:	Update a specific description (using keyword "all")
Endpoint:	{HOST}/descriptions/all/{description_id}
HTTP Method:	PUT
Description:	With this endpoint, the providers of the descriptions and the administrators, are able to update the content of the (/their) descriptions. It is possible to update the whole description or only some fields. It requires the ID of the description to be at the end of the endpoint and the users' API key in the headers of the PUT request. As in the create action, the description should be provided as raw data in JSON format. It should be noted that this endpoint uses the keyword "all" (the descriptions are already stored in the marketplace, thus the platform knows the collections in which they have been stored). Below is the standard schema for a description:

```
{
   "title": "<title of the asset>",
   "description ": "<description of the provided
asset>",
   "type": "<type of the asset (same as the
collection value)>",
   "owner ": "<organization / author / etc.>",
   "contact": "<provider's name and email>",
   "availability": "<reflects the users who can
see the description:          public / infinitech
/ specific Work Packages or Pilots / Other>",
"keywords": ["<keyword 1>", … ], "comments":
"<provider's comments>"
}
```

Body Data:	Raw (JSON) Data – as the above or similar schema (Content-Type: application/json). The descriptions can also be updated from binary files that contain the above JSON schema (curl example can be found at the end of the interface).	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	description_id	The ID of the description that will be updated.
Query Parameters:	None	
Restrictions/Special Features:	Available only for the providers/creators of the descriptions and for the administrators who can update any description.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request PUT '{HOST}/descriptions/all/{description_id}' \
--header 'API_KEY: <value>' --header 'Content-Type:
application/json' \ --data-raw '{
   "existing_or_new_field": "<value>", …
}'
```

Table 2.27. Interface: Update a specific description (using description's "collection").

Title:	Update a specific description (using description's "collection")	
Endpoint:	{HOST}/descriptions/{collection}/{description_id}	
HTTP Method:	PUT	
Description:	This PUT request is similar to the previous request. The only difference is that instead of using keyword "all" it uses a specific collection/the collection in which the description that will be updated has been stored during its creation/initial upload. The endpoint is restricted and available only to descriptions' providers/creators and to administrators who can update any description. Thus, the API keys of the requesters must be included in the headers of the request. More information about the endpoint can be found on the previous endpoint.	
Body Data:	Raw (JSON) Data – as the schema of the previous endpoint (Content-Type: application/json). It should be noted that the descriptions can also be updated from binary files that contain the JSON schema of the previous endpoint (example in curl can be found at the end of the interface).	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
	description_id	The ID of the description that will be updated.
Query Parameters:	None	
Restrictions/Special Features:	Available only for the providers/creators of the descriptions and for the administrators who can update any description.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request PUT
'{HOST}/descriptions/{collection}/{description_id}'
\ --header 'API_KEY: <value>' --header 'Content-
Type: application/json' \ --data-raw '{
   "existing_or_new_field": "<value>", …
}'
```

Table 2.28. Interface: Delete a specific description (using keyword "all").

Title:	Delete a specific description (using keyword "all")
Endpoint:	{HOST}/descriptions/all/{description_id}
HTTP Method:	DELETE

(*Continued*)

Table 2.28. Continued

Description:	A DELETE request to this endpoint has as a result the deletion of a specific description, using its ID. The endpoint is restricted and available only to descriptions' providers/creators and to administrators who can delete any description. Thus, users' API keys must be included in the headers of the request. It should be noted that this endpoint uses the keyword "all" instead of description's collection (the descriptions are already stored in the Marketplace, thus the platform knows the collections in which have been stored).	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	description_id	The ID of the description that will be deleted.
Query Parameters:	None	
Restrictions/Special Features:	Available only for the providers/creators of the descriptions and for the administrators who can delete any description.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/descriptions/all/{description_id}' --
header 'API_KEY: <value>'
```

Table 2.29. Interface: Delete a specific description (using description's "collection").

Title:	Delete a specific description (using description's "collection")	
Endpoint:	{HOST}/descriptions/{collection}/{description_id}	
HTTP Method:	DELETE	
Description:	This DELETE request is similar to the above request, with the difference that, instead of keyword "all", it uses a specific collection/the collection in which the description that will be deleted has been stored during its creation/initial upload. The endpoint is restricted and available only to descriptions' providers/creators and to administrators who can delete any description. Thus, the API keys of the requesters must be included in the headers of the request. More information about the endpoint can be found on the previous endpoint.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "third-party-tools", "whitepapers", "how-to-videos", "accelerator-programmes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
	description_id	The ID of the description that will be deleted.

(*Continued*)

Table 2.29. Continued

Query Parameters:	None
Restrictions/Special Features:	Available only for the providers/creators of the descriptions and for the administrators who can delete any description.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request DELETE
'{HOST}/descriptions/{collection}/{description_id}' \
--header 'API_KEY: <value>'
```

Table 2.30. Interface: Delete all descriptions.

Title:	Delete all descriptions
Endpoint:	{HOST}/descriptions/all/all
HTTP Method:	DELETE
Description:	This endpoint is available only to the administrators, who can delete all the existing descriptions from all the collections (the keyword "all" is used instead of a specific collection). The endpoint is restricted and thus, users' API keys must be included in the headers of the request. For security reasons, the requesters should provide their password in the body of their request, as raw data (JSON schema): `{ "password": "…" }`
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json).
Headers:	Key: API_KEY; Value: Requester's API key.
URL Parameters:	Parameter: None; Value: None
Query Parameters:	None
Restrictions/Special Features:	Available only for the administrators of the Marketplace.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/descriptions/all/all' \
--header 'API_KEY: <value>' --header 'Content-Type: application/json'
\
--data-raw '{ "password": "…" }'
```

Table 2.31. Interface: Delete all descriptions from a specific collection.

Title:	Delete all descriptions from a specific collection
Endpoint:	{HOST}/descriptions/{collection}/all
HTTP Method:	DELETE

(*Continued*)

Table 2.31. Continued

Description:	This DELETE request is similar to the above request and it is only available to administrators who can delete all the descriptions from a specific collection. It is a restricted endpoint and thus, users' API keys must be included in the headers of the request. For security reasons, the requesters should provide their password in the body of their request, as raw data (JSON schema): `{ "password": "…" }`
Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json).
Headers:	Key: API_KEY Value: Requester's API key.
URL Parameters:	Parameter: collection Value: Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "third-party-tools", "whitepapers", "how-to-videos", "accelerator-programmes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
Query Parameters:	None
Restrictions/Special Features:	Available only for the administrators of the Marketplace.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/descriptions/{collection}/all' \
--header 'API_KEY: <value>' --header 'Content-Type: application/json'
\
--data-raw '{ "password": "…" }'
```

Table 2.32. Interface: Get a list with all descriptions that need permission (administrators' action).

Title:	Get a list with all descriptions that need permission (administrators' action)
Endpoint:	{HOST}/descriptions/permit/all
HTTP Method:	GET
Description:	This endpoint returns the descriptions from all the collections (since the keyword "all" is used) that need permission before they become available to the Marketplace's users. A description needs permission either when it is uploaded or after it has been updated by the users. Moreover, the endpoint is only available to administrators and thus, the API key of a requester is required in the headers of the request. Finally, the endpoint offers some standard query parameters that specify the format of the results and are described below (Query Parameters).
Body Data:	None
Headers:	Key: API_KEY Value: Requester's API key.
URL Parameters:	Parameter: None Value: None

(*Continued*)

Table 2.32. Continued

Query Parameters:	Key	Value
	sortBy	[Optional] Sorts the descriptions by a field – the default is the "newest" key. The value should be one of the following: "**newest**": sort by date in descending order. "**oldest**": sort by date in ascending order. "**title**": sort by title in ascending order.
	itemsPerPage	[Optional] Returns the results separated in pages (arrays) of N items. The number N is specified by the value of this key. The value N must be an integer number greater or equal to 1. If the key is not used or has a non-accepted value, the results are returned on a single page.
	Page	[Optional] This key can only be used if the "itemsPerPage" key is also used. If it is used, it returns the specified (by key's value) page instead of all pages created using the key "itemsPerPage". The value must be an integer greater or equal to 1. The default value is 0, meaning that all pages will be returned.
Restrictions/Special Features:	Available only to the administrators.	
Successful Response:	JSON Object with the descriptions (from all collections) that need permission in its content.	

The following is an example of the request in cURL:

```
+ curl -request GET '{HOST}/descriptions/permit/all' --header 'API_KEY:
<value>'
+ curl -request GET '…?sortBy={value}' …
+ curl -request GET '…?itemsPerPage={value}' …
+ curl -request GET '…?itemsPerPage={value}&page={value}' …
+ curl -request GET '…?sortBy={value}&itemsPerPage={value}' …
+ curl -request GET '…?sortBy={value}&itemsPerPage={value}&page={value}' …
```

Table 2.33. Interface: Get a list with all descriptions from a specific collection that need permission (administrators' action).

Title:	Get a list with all descriptions from a specific collection that need permission (administrators' action)	
Endpoint:	{HOST}/descriptions/permit/{collection}	
HTTP Method:	GET	
Description:	This request is similar to the above request. The only difference between the two actions is that the current request retrieves the descriptions that need permission from a specific collection (uses a specific {collection} value instead of the keyword "all"). For more details, refer to the above endpoint.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.

(*Continued*)

Table 2.33. Continued

URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
Query Parameters:	As in the above request.	
Restrictions/Special Features:	Available only to the administrators.	
Successful Response:	JSON Object with the descriptions (from a specific collection) that need permission in its content.	

The following is an example of the request in cURL:

```
+ curl -request GET '{HOST}/descriptions/permit/all' --header 'API_KEY:
<value>'
** The previous endpoint's examples also appry to the current one, but
using a {collection} value
```

Table 2.34. Interface: Approve or reject a description that needs permission, using keyword "all" (administrators' action).

Title:	Approve or reject a description that needs permission, using keyword "all" (administrators' action)	
Endpoint:	{HOST}/descriptions/permit/all/{description_id}	
HTTP Method:	POST	
Description:	This endpoint is used by administrators to approve or reject a specific description (using its ID) that needs administrators' permission. The endpoint is restricted and available only to administrators and thus, the requesters must provide their API key in the headers of the request. Also, this endpoint uses the keyword "all" and not the collection in which a specific description is stored, as the next endpoint does. An important parameter/key that must be included in the request's headers is the "xpermission" key that should contain the word "approve" for the description's approval, otherwise the word "disapprove" for its rejection. A rejection/disapproval of a description results to the deletion of the description and all its assets/contents.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-permission	Valid values: {"approve", "disapprove"}
URL Parameters:	Parameter	Value
	description_id	The ID of the description that will be approved or rejected.
Query Parameters:	None	
Restrictions/Special Features:	Available only to the administrators.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl -request POST '{HOST}/descriptions/permit/all/{description_id}' \
--header 'API_KEY: <value>' --header 'x-permission: <value>'
```

Table 2.35. Interface: Approve or reject a description that needs permission, using description's "collection" (administrators' action).

Title:	Approve or reject a description that needs permission, using description's "collection" (administrators' action)	
Endpoint:	{HOST}/descriptions/permit/{collection}/{description_id}	
HTTP Method:	POST	
Description:	This request is similar to the above request. The only difference between the two actions is that the current request uses the value of the {collection} in which a specific descriptions is stored. For more details, refer to the above endpoint.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-permission	Valid values: {"approve", "disapprove"}
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses ", "containers", "uncategorized", "other"}
	description_id	The ID of the description that will be approved or rejected.
Query Parameters:	None	
Restrictions/Special Features:	Available only to the administrators.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl -request POST
'{HOST}/descriptions/permit/{collection}/{description_id}' \
--header 'API_KEY: <value>' --header 'x-permission: <value>'
```

Table 2.36. Interface: Approve or reject all descriptions that need permission, using keyword "all" (administrators' action).

Title:	Approve or reject all descriptions that need permission, using keyword "all" (administrators' action)
Endpoint:	{HOST}/descriptions/permit/all/all
HTTP Method:	POST
Description:	This endpoint is used by the administrators to approve or reject all the stored descriptions (from all the collections, since keyword "all" is used) that need administrators' permission. The endpoint is restricted and available only to administrators and thus, the requesters' must provide their API keys in the headers of the request. An important parameter/key that must be included in the headers of the request is the "x-permission" key that should have as a value the word "approve" for the descriptions to be approved, otherwise the word "disapprove" to be rejected. A rejection/disapproval of the descriptions has as a result the deletion of the descriptions and all of their assets and contents.

(*Continued*)

Table 2.36. Continued

Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-permission	Valid values: {"approve", "disapprove"}
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	None	
Restrictions/Special Features:	Available only to the administrators.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl -request POST
'{HOST}/descriptions/permit/all/all' \ --
header 'API_KEY: <value>' --header 'x-
permission: <value>'
```

Table 2.37. Interface: Approve or reject all descriptions that need permission under a specific collection, using a "collection" value (administrators' action).

Title:	Approve or reject all descriptions that need permission under a specific collection, using a "collection" value (administrators' action)	
Endpoint:	{HOST}/descriptions/permit/{collection}/all	
HTTP Method:	POST	
Description:	This request is similar to the above request. The only difference is that the administrators, using the current endpoint, are able to approve or reject all the descriptions of a specific {collection}. The endpoint is restricted and available only to administrators and thus, the requesters' must provide their API keys in the headers of the request. An important parameter/key that must be included in the headers of the request is the "x-permission" key that should have as a value the word "approve" for the descriptions to be approved, otherwise the word "disapprove" to be rejected. A rejection/disapproval of the descriptions has as a result the deletion of the descriptions and all of their assets and contents.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-permission	Valid values: {"approve", "disapprove"}
URL Parameters:	Parameter	Value
	collection	Valid values: {"algorithms", "notebooks", "ml-models", "datasets", "apis", "webinars", "applications", "blockchain", "thirdparty-tools", "whitepapers", "how-to-videos", "acceleratorprogrammes", "innovation-support-services", "workshops", "courses", "containers", "uncategorized", "other"}

(*Continued*)

Table 2.37. Continued

Query Parameters:	None
Restrictions/Special Features:	Available only to the administrators.
Successful Response:	JSON Object with a successful message.

The following is an example of the request in cURL:

```
curl -request POST '{HOST}/descriptions/permit/{collection}/all' \
--header 'API_KEY: <value>' --header 'x-permission: <value>'
```

2.5 Search Functionality on Descriptions

The search functionality is a vital requirement for most services in order to reduce the number of objects returned by a query. Thus, the back-end's endpoints that retrieve multiple descriptions simultaneously, support some relative query filters. These filters enable the users of the marketplace to search for assets, based on various parameters from the content of the stored descriptions.

More specifically, the interfaces of the back-end that return lists of assets, support additional query parameters with any key-value pair. Query parameters are a defined set of parameters attached to the end of a URL and are used in order to help search specific content or actions based on the data being passed. In order to append query parameters to the end of a URL, a question mark "**?**" is added to the end of the URL, followed immediately by a pair of a key and a value, separated by an equal symbol "**=**". Moreover, a URL can have multiple parameters, by adding an ampersand symbol "**&**" between each pair of key-value.

In the context of the INFINITECH Marketplace and the description, the keys added to the URLs as query parameters must be valid, in the sense that they exist as fields in the descriptions and their search has a real value. Below are some valid syntaxes for advanced search with additional query parameters. The examples use the "Get a list with all descriptions" interface.

```
Single attribute:
'{HOST}/descriptions/all?<attribute_name>=<value>'
More attributes:
'{HOST}/descriptions/all?<attribute_1>=<value>&<attribute_2>=<va
lue>&…'
```

Moreover, the Python programming language that is used by the back-end (as described in Section 3.1), enables access to nested fields of dictionary/JSON object using a dot "**.**" between a key at the first level of the hierarchy and a key at the second level (this applies to all levels, up to the lowest level). Thus, the next example is also a valid schema of a query:

```
For attributes in lower hierarchical level: "
'{HOST}/descriptions/all?<attribute_level_1>.<…>.<attribute_level
_n>=<value>'
```

To sum up, given the above syntaxes of a valid query, the following search example request in cURL, returns the descriptions that in their title contain the value "machine learning" and their type is "algorithms":

```
curl                          --request                          GET
'{HOST}/descriptions/all?title=machine%20learning&type=algorithm'
```

It should be noted that the value "%20" is the ASCII Encoding Reference of the space character.

Except for these, the back-end supports advanced searching using some operators that extend the keys of the query parameters, using a dot "." between the keys and the operators. Below are the supported operators along with a description for their usage.

Below are some examples of using the previous operators.

```
eq: '{HOST}/descriptions/all?metadata.provider.eq=vkoukos' ne:
'{HOST}/descriptions/all?metadata.version.ne=1' gt:
'{HOST}/descriptions/all?metadata.views.gt=100' gte:
'{HOST}/descriptions/all?info.type.gte=datasets' lt:
'{HOST}/descriptions/all?metadata.uploadDate.lt=2022-10-15' lte:
'{HOST}/descriptions/all?metadata.views.lte=20'
in:
'{HOST}/descriptions/all?info.title.in=machine,learning,algorithm'
nin: '{HOST}/descriptions/all?info.keywords.nin=economy,finance'
```

Furthermore, the back-end's search mechanism uses a ranking system for the results. More specifically, for each description in the results, it maintains a score resulting from the points it receives for each search argument.

In an equality search (using "=" symbol or "eq" operator) for a specific key, the points that a description receives can be one of the following:

- **5**: if the values are exactly equal (same) and case sensitive.
- **4**: if the values are equal (same) but not case sensitive.
- **3**: if the values are similar (e.g., the first value contains the second value but are not the same) and case sensitive.
- **2**: if the values are similar but not case sensitive.
- **0**: if the values do not match.

The other operators just receive **1** point if the conditions match (i.e. the condition is "true"). The operator "in" uses the operator "eq" (or the symbol "=") for each value in its "array" and thus, it has the same score system. Finally, the operator "nin" uses the operator "ne" for each value in its "array" (Table 2.38).

Table 2.38. Back-end's search operators.

Operator	Usage	Example
eq	Full title: **equal** This operator performs an equality search and has exactly the same use with the equality symbol "=". It applies to both texts (strings) and numbers.	<key>.eq=<value>
ne	Full title: **not equal** This operator performs a non-equality search. It applies to both texts (strings) and numbers.	<key>.ne=<value>
gt	Full title: **greater than** This operator performs searching for a key with a value greater than the provided. It applies to both texts (strings) and numbers.	<key>.gt=<value>
gte	Full title: **greater than or equal** This operator performs searching for a key with a value greater than or equal to the provided. It applies to both texts (strings) and numbers.	<key>.gte=<value>
lt	Full title: **less than** This operator performs searching for a key with a value less than to the provided. It applies to both texts (strings) and numbers.	<key>.lt=<value>
lte	Full title: **less than or equal** This operator performs searching for a key with a value less than or equal to the provided. It applies to both texts (strings) and numbers.	<key>.lte=<value>
in	Full title: **in (equal to one of the values)** This operator performs searching for a key with a value equal to one of the provided values. The <value> may have multiple values separated by a comma ",". It applies to both texts (strings) and numbers.	<key>.in=<value_1>,<value_2>
nin	Full title: **not in (not equal to any of the values)** This operator performs searching for a key with a value not equal to any of the provided values. The <value> may have multiple values separated by a comma ",". It applies to both texts (strings) and numbers.	<key>.nin=<value_1>,<value_2>

2.6 APIs Related to Assets

This group of APIs offers functionalities intended for the management of the assets. They support all CRUD operations for files which are stored in the marketplace. Table 2.39 presents the endpoints related to Assets as they are in the last version of the marketplace's back-end.

- {*HOST*} refers to the hosting server: the domain name and the port that the back-end runs.
- {asset_*id*} refers to the ID of a specific asset.
- {*description_id*} refers to the ID of the description with which the new asset will be linked to.
- {*given_asset_id*} is used in "upload a new asset" action, providing new-asset's ID.

Table 2.39. APIs of the back-end related to Assets.

Action	HTTP Method	Endpoint
Get a list with the stored assets	GET	{HOST}/assets
Get a specific asset, using its ID	GET	{HOST}/assets/{asset_id}
Upload a new asset with random ID, linked to a specific description	POST	{HOST}/assets/{description_id}
Upload a new asset with given ID, linked to a specific description	POST	{HOST}/assets/{description_id}/{given_asset_id}
Update a specific asset, using its ID	PUT	{HOST}/assets/{asset_id}
Delete a specific asset, using its ID	DELETE	{HOST}/assets/{asset_id}
Delete all assets (administrators' action)	DELETE	{HOST}/assets/all

Table 2.40. Interface: Get a list with the stored assets.

Title:	Get a list with the stored assets	
Endpoint:	{HOST}/assets	
HTTP Method:	GET	
Description:	A GET request to this endpoint will result in the retrieval of a list with the stored assets and some additional information of them. The endpoint is restricted and available only to administrators and thus, the requesters' must provide their API keys in the headers of the request.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	None	
Restrictions/Special Features:	Available only to administrators.	
Successful Response:	Results in JSON Object.	

The following is an example of the request in cURL:

```
curl --request GET '{HOST}/assets' --header 'API_KEY: <value>'
```

- Some of these actions require additional fields in the headers or even in the body of the HTTP request. Example of a required field is the API key that users use in order to validate themselves to the platform.

Table 2.40 to 2.46 provide details on each of these actions respectively.

Table 2.41. Interface: Get a specific asset, using its ID.

Title:	Get a specific asset, using its ID
Endpoint:	{HOST}/assets/{asset_id}
HTTP Method:	GET

(*Continued*)

Table 2.41. Continued

Description:	This endpoint is used to retrieve a specific stored asset. For the retrieval of the assets, the usage of the asset's ID is necessary. Also, this endpoint is restricted and thus, the API key of a requester must be included in the headers of the request.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	asset_id	The ID of the asset that will be retrieved.
Query Parameters:	None	
Restrictions/Special Features:	Available to all authenticated (and verified) users.	
Successful Response:	Binary data	

The following is an example of the request in cURL:

```
curl --request GET '{HOST}/assets/{asset_id}' --header 'API_KEY:
<value>'
```

Table 2.42. Interface: Upload a new asset with random ID, linked to a specific description.

Title:	Upload a new asset with random ID, linked to a specific description	
Endpoint:	{HOST}/assets/{description_id}	
HTTP Method:	POST	
Description:	Through this endpoint the users can upload their assets. It requires to add at the end of the endpoint the description's ID with which is going to be linked. It is also needed to add to the headers of the request the API key of the provider and the asset's filename, whilst the assets must be uploaded as form-data with the key "asset".	
Body Data:	Data Type:	Form Data
	Key	Value
	asset	Binary data/Path to file
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-filename	New asset's filename.
URL Parameters:	Parameter	Value
	description_id	The ID of the description with which the new asset is going to be linked.
Query Parameters:	None	
Restrictions/Special Features:	Available for the descriptions' providers with which the assets will be linked, and also for the administrators who can upload assets on behalf of the providers.	
Successful Response:	JSON Object with the new asset's ID in its content.	

The following is an example of the request in cURL:

```
curl --request POST
'{HOST}/assets/{description_id}' \ -
-header 'API_KEY: <value>' --header
'x-filename: <value>' \
--form 'asset=@"<full_path_to_asset>"'
```

Table 2.43. Interface: Upload a new asset with given ID, linked to a specific description.

Title:	Upload a new asset with given ID, linked to a specific description	
Endpoint:	{HOST}/assets/{description_id}/{given_asset_id}	
HTTP Method:	POST	
Description:	This endpoint is similar to the previous. The difference is that with the current endpoint it is possible to specify the ID of the new asset, providing it at the end of the endpoint {given_asset_id}. Currently, this endpoint can be used only by the administrators of the market platform.	
Body Data:	Data Type:	Form Data
	Key	Value
	asset	Binary data/Path to file
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-filename	New asset's filename.
URL Parameters:	Parameter	Value
	description_id	The ID of the description with which the new asset is going to be linked.
	given_asset_id	The ID to be given to the new asset.
Query Parameters:	None	
Restrictions/Special Features:	Available only for administrators whether they upload an asset for their descriptions or upload an asset on behalf of the providers.	
Successful Response:	JSON Object with the new asset's ID in its content.	

The following is an example of the request in cURL:

```
curl --request POST '{HOST}/assets/{description_id}/{given_asset_id}' \
--header 'API_KEY: <value>' --header 'x-filename: <value>' \
--form 'asset=@"<full_path_to_asset>"'
```

Table 2.44. Interface: Update a specific asset, using its ID.

Title:	Update a specific asset, using its ID	
Endpoint:	{HOST}/assets/{asset_id}	
HTTP Method:	PUT	
Description:	With this PUT request, it is possible to update an already stored asset. The asset's ID that should be included at the end of the endpoint, determines which asset should be replaced by the newly provided asset. As in the uploading, the asset should be uploaded as form-data with the key "asset" and the headers of the request have to contain provider's API key. It should be noted that the users can only update the assets provided by themselves (except for administrators).	
Body Data:	Data Type:	Form Data
	Key	Value
	asset	Binary data/Path to file
Headers:	Key	Value
	API_KEY	Requester's API key.
	x-filename	[Optional] Asset's new filename.

(*Continued*)

Table 2.44. Continued

URL Parameters:	Parameter	Value
	asset_id	The ID of the asset that will be updated.
Query Parameters:	None	
Restrictions/Special Features:	Available only for the providers of the descriptions/assets and for the administrators who can update stored assets on behalf of the providers.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request PUT '{HOST}/assets/{asset_id}' \
--header 'API_KEY: <value>' --header 'x-filename: <value>' \
--form 'asset=@"<full_path_to_asset>"'
```

Table 2.45. Interface: Delete a specific asset, using its ID.

Title:	Delete a specific asset, using its ID	
Endpoint:	{HOST}/assets/{asset_id}	
HTTP Method:	DELETE	
Description:	A DELETE request to this endpoint has as a result the deletion of a specific asset, by using its ID in order to find it. This endpoint is restricted and thus, users' API keys must be included in the headers of the request. It should be noted that an asset can be deleted only by its provider and the administrators.	
Body Data:	None	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	asset_id	The ID of the asset that will be deleted.
Query Parameters:	None	
Restrictions/Special Features:	Available only for the providers of the descriptions/assets and for the administrators who can delete any stored assets.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/assets/{asset_id}' --header 'API_KEY:
<value>'
```

Table 2.46. Interface: Delete all assets (administrators' action).

Title:	Delete all assets (administrators' action)
Endpoint:	{HOST}/assets/all
HTTP Method:	DELETE
Description:	This DELETE request is similar to the above request, with the difference that it deletes all the assets, as it uses the keyword "all". Again, it is necessary the usage of the users' API keys and **it is only available to administrators**. For security reasons, the requesters should provide their password in the body of their request, as raw data (JSON schema): `{ "password": "…" }'`

(*Continued*)

Table 2.46. Continued

Body Data:	Raw (JSON) Data – as the above schema (Content-Type: application/json).	
Headers:	Key	Value
	API_KEY	Requester's API key.
URL Parameters:	Parameter	Value
	None _id	None
Query Parameters:	None	
Restrictions/Special Features:	Available only to the administrators of the platform.	
Successful Response:	JSON Object with a successful message.	

The following is an example of the request in cURL:

```
curl --request DELETE '{HOST}/assets/all' --header
'API_KEY: <value>' \  --header 'Content-Type:
application/json' --data-raw '{ "password": "…" }'
```

2.7 Other Interfaces

One endpoint that was not mentioned is that of the back-end's root interface, which presents a roadmap of the main back-end's interfaces. The latter is described at the Table 2.47:

Table 2.47. Interface: Root interface.

Title:	Root interface	
Endpoint:	{HOST}	
HTTP Method:	GET	
Description:	This endpoint returns a list with the back-end's interfaces that are available to be used by all users. It acts as a roadmap, providing the interfaces along with short information about the functionalities that they trigger. The structure of the information follows a tree approach.	
Body Data:	None	
Headers:	Key	Value
	None	None
URL Parameters:	Parameter	Value
	None	None
Query Parameters:	None	
Restrictions/Special Features:	None	
Successful Response:	Back-end's roadmap in text/plain.	

The following is an example of the request in cURL:

```
curl --request GET '{HOST}'
```

Except for the already described interfaces, the back-end provides the following restricted interfaces that will not be described, since they are mostly related only

with the platform's administrators. In short, the interfaces' titles are:

- Get a list of the administrators,
- Add a new administrator,
- Remove an administrator,
- Get system's backup,
- Restore a backup,
- Get a list of all users,
- Get a list with system metrics/report.

2.8 Validation Scenarios

This section presents some validation scenarios regarding the functionalities of the back-end component of the marketplace in combination with the endpoints described in the previous sections.

Specifically, the scenarios that will be presented are how users can be authenticated to the back-end, how they can upload a new description, how they can upload an asset for the description they created in the second scenario and how users can retrieve an asset that matches their interests.

It is noted that all scenarios will be done exclusively using the back-end. Similar scenarios, but using the frontend, are presented in deliverable D8.6 "IoT and Blockchain Solutions Marketplace – II".

2.8.1 Authentication Scenario

We consider a registered user in the INFINITECH Marketplace with username "test_user" and password "test_password". The user, as described in Section 3.2.1, in order to be authenticated, needs to send a POST HTTP request to the endpoint *"{HOST}/accounts/users/authentication"*. In the body of the request, the user should add its credentials in the following format:

```
{ "username": "test_user", "password": "test_password" }
```

Thus, according to the description of the endpoint, the user should send the following request (using cURL):

```
curl --request POST
'{HOST}/accounts/users/authentication' \ --header
'Content-Type: application/json' \
--data-raw '{ "username": " test_user", "password": "
test_password" }'
```

The response of this request, if everything is fine (e.g. correct credentials), will be a temporary code that acts as an API key. Through this API key, the user is able to use the rest functionalities of the marketplace.

2.8.2 Upload Description Scenario

In this scenario, we will use the user of the previous scenario and we consider that the user followed the above scenario in order to be authenticated to the marketplace (i.e. has an API key).

The same user wants to upload a new asset/solution to the Marketplace. In this case, the first step that need to do, is to upload/create a new description to the Marketplace.

All descriptions have a specific schema and should be filled in accordingly. Using the front-end this is done automatically through the corresponding forms, but in case the request is made in the back-end, the description must be created by the user.

As described in Section 3.2.2, the user must add this description in JSON format as raw data to the body of the request. We suppose that the user wants to upload the description for the already uploaded description for the Python Notebook "DeepVaR" (https://marketplace.infinitech-h2020.eu/infi_assets/deepvar-valueat-risk-prediction-leveraging-deep-learning). The user – provider prepares the corresponding description as follows:

```
{
   "title": "DeepVaR: Value-at-Risk prediction leveraging Deep
Learning",
   "description ": "This component both explains and implements the
DeepVaR, which is a Value-at-    Risk model based on deep neural
networks and Monte Carlo simulations.",
   "type": "Notebook", "owner ": "INNOV-ACTS LTD", "contact": "…",
"availability": "INFINITECH",
   "keywords": [ "Finance","Risk
Management" ], "comments": null }
```

In order to upload this description to the marketplace, the user should send a POST request to the endpoint *"{HOST}/descriptions/{collection}"* where the {*collection*} will be the value "notebooks". Also, the user should include the API key that got as a response during the authentication (previous scenario).

The following is an example of the corresponding request in cURL:

```
curl --request POST '{HOST}/descriptions/notebooks' \
--header 'API_KEY: <value_retrieved_from_scenario_1>' --header
'Content-Type: application/json' \ --data-raw '{
   "title": "DeepVaR: Value-at-Risk prediction leveraging Deep
Learning",
   "description ": "This component both explains and implements the
DeepVaR, which is a Value-at-    Risk model based on deep neural
networks and Monte Carlo simulations.",
   "type": "Notebook", "owner ": "INNOV-ACTS LTD", "contact": "…",
"availability": "INFINITECH",
   "keywords": [ "Finance","Risk
Management" ], "comments": null }'
```

The response of this request will be the ID of the new description. This ID will be used in the next scenario, in order to upload the description's asset.

2.8.3 Upload Asset Scenario

This scenario extends the previous scenarios. It uses the API key of the user that has been authenticated as per the first scenario, and it uses the ID of the description that uploaded in the second scenario. In this scenario, the same user wants to upload an asset for the description of DeepVaR Notebook. The user has a compressed file (.zip) to upload which need to be linked with a specific description.

Thus, as described in Section 3.2.4, the user should send a POST request to the next endpoint *"{HOST}/assets/{description_id}"* where the "*description_id*" is the ID that got as a response in the second scenario, during the uploading of the description to which the asset will be linked. Moreover, the user should add to the headers of the request a) the API key from the first scenario and b) the asset's filename (e.g. deepvar.zip).

The following is an example of the scenario's request in cURL:

```
curl --request POST '{HOST}/assets/{description_id}' \
--header 'API_KEY: <value_retrieved_from_scenario_1>' --header 'x-
filename: deepvar.zip' \
--form 'asset=@"./deepvar.zip"'
```

The response of the request is the new asset's ID and the result, except for the storing of the asset, is the connection between the asset and the description.

2.8.4 Retrieve Asset Scenario

This scenario combines various interfaces, even the authentication interface that has been used in the first scenario, to describe how to retrieve an asset from a description that the user found after a related search in the stored descriptions/solutions of the Marketplace that cover user's specific needs.

We consider an INFINITECH Marketplace user who wants to find a deep learning model for identifying Value at Risk occasions (financial metric that estimates the risk of an investment). So, the first step that user should do is to search for descriptions that are relevant to this user's need. Based on the Sections 3.2.2 (APIs related to Descriptions) and 3.2.3 (Search functionality on Descriptions), the user should send a GET request to the next endpoint *"{HOST}/descriptions/all"* (i.e. search on all descriptions & collections – Table 2.16) using additional query parameters with such key-value pairs to retrieve the descriptions that match to the "Value at Risk" need.

More specifically, the user will add as query parameters the following key-value pairs and operators:

Key	Operator	Value
title	in	value, risk
keywords	"eq" or just "="	finance
description	in	Value-at-Risk, model

The following is the request (in cURL) for user's first step (i.e. searching):

```
curl --request GET '{HOST}/descriptions/all?info.title.in=value,risk
                                            &info.keywords=finance
                                            &info.description.in=Value-
at-Risk,model"'
```

The response of this request is a list of descriptions in JSON format. Moreover, as explained in Table 2.16, this interface returns the descriptions in their short schema. Thus, as the user views the results, finds an interesting description that may meet the desired needs. The description that the user found is the DeepVaR Notebook that has been used in the second scenario and its short description schema is the following:

```
{
 "collection": "notebooks", "id": "notebooks_agWabSJaFsgse",

  "info": { "title": " DeepVaR: Value-at-Risk prediction leveraging
Deep Learning.",
      "keywords": [ "Finance","Risk Management" ], "owner": "INNOV-
ACTS LTD",

      "short_desc": "This component both explains and implements the
DeepVaR, which is a Value-at-

                     Risk model based…", "type": " notebook"
  }, metadata": {"provider": "…","updateDate": "…",
 "uploadDate": "…", "views": 135 }

}
```

From the above description, the user retrieves the "ID" of the description (i.e. "*notebooks_agWabSJaFsgse*") in order to use it in the second step which is to retrieve the full schema (and information) of this solution/description.

This step can be done by making a GET request to Table's 2.18 interface "*Get a specific description (using keyword "all")*", i.e. to *"{HOST}/descriptions/all/{description_id}"*, where the {*description_id*} will be replaced by the ID of the above description. The user can also use the interface of Table 2.19, where the collection of the description should be specified, which is known since the collection of the description is included in the short schema of the description (i.e. "notebooks").

It should be noted that the user should use an API key (retrieved after authentication – scenario 1) in order to retrieve the full schema of the description. The following is the request (in cURL) for user's second step (i.e. description's retrieval):

```
curl --request GET '{HOST}/descriptions/all/notebooks_agWabSJaFsgse' \
     --header 'API_KEY: <value>'
```

The response of this request is a JSON Object with the description's full schema in the results. Following the scenario 3, where an asset for the "DeepVaR" has been uploaded, the full schema of the description, is the following:


```
{
   "id": "notebooks_agWabSJaFsgse",
   "collection": "notebooks",
   "info": {
            "title": "DeepVaR: Value-at-Risk prediction leveraging
Deep Learning.",
            "type": "notebook",
            "comments": null,
            "contact": "…",
            "description": "This component both explains and
implements the DeepVaR, which is               a Value-at-Risk
model based on deep neural networks and Monte Carlo
simulations.",
            "keywords": [ "Finance", "Risk
Management" ],              "owner": "INNOV-ACTS
LTD", "availability": "INFINITECH", "license": "-",
            "subtype": null, "fieldsOfUse": []
   },
   "metadata": {
                "approved": 1, "last_updated_by": "…", "md5": "…",
                "provider": "…", updateDate": "…", "uploadDate":
"…",
                "version": 2, "views": 135},
   "assets": [
          {
            "verified": 1, "downloads": 40, "filename":
"deepvar.zip",
            "id": "80F7MjRTIxvb-7qIKRAjv-IJ3p-b3vL", "version": 1,
            "md5": "…", "size": "…", "updateDate": "…"
          }, {…}, {…}, … ,{…} !
    ],
    "resources": []
}
```


The last step for the user in order to retrieve an asset that seems helpful, covering user's needs, is to send a GET request to the following endpoint "*{HOST}/assets/{asset_id}*" where the {*asset_id*} is the ID of description's asset which

can be found in the "*id*" field under the "assets" field of the above description's full schema. As described in Section 3.2.4 and in Table 2.41, the endpoint is restricted and thus, the user needs also to provide an API key (scenario 1) in the headers of the request.

Thus, the user should make the following request (example in cURL):

```
curl --request GET '{HOST}/assets/80F7MjRTIxvb-7qIKRAjv-IJ3p-b3vL' \
     --header 'API_KEY: <value>'
```

This request returns the binary data (as a file/depending on the type of asset) of the specified-by-its-ID asset which can be used by the user for work, scientific or other reasons.

DOI: 10.1561/9781638282334.ch3

Chapter 3

INFINITECH Technology Assets

3.1 Anti-Money Laundering

Spyros X

Fraud detection & money laundering detection system. The fraud detection system is ML-based, while the money laundering system is Rule Based. It is Business to Business model and it is for NFT, Crypto currency and Bank owners that want to secure their networks.

Graph-based Anomaly Data Platform

Webtool including a demo to test a variety of use-cases (e.g. self financing by bank account/recharge, etc.). It allows both loading in-built Nexi uses cases on synthetic data and importing external files.

Graph-based Anomaly Exploration Platform

Graph-based Anomaly Exploration Platform as Jupyter Server engine for data scientists. It is a server based Jupyter notebook which allows expert data professionals to develop scripts and perform a more customized analysis. The data specialist can interact with algorithm outputs or payments data: it is possible for him/her to write code and therefore explore results with customized queries and/or algorithms.

3.2 API

Cross Over Open Banking Hub v1.0

CrossOver is a PSD2 engine that helps developers and fintechs make use of all PSD2 APIs using a single interface. It handles the entire lifecycle of consents, stores the users' accounts, and executes payments so you can focus on the core aspects of your application.

Reportbrain News Sentiment API

Reportbrain News API – Facilitate News features using our structuring of unstructured data in the News by transforming, normalizing, and augmenting it for a specific application. Customers can use Reportbrain's query language to explore that data and retrieve news articles and their metadata. The Reportbrain News API is accessed through HTTP via a list of available calls and their parameters.

INFINITECH Graph Data Model Online Tool

Data Interoperability for Fintech's and Financial Sector. This online tool refers to the INFINITECH Project Online Ontology Mapping Framework and Toolkit, it includes the Graph Data Model, the Data Sharing Files and Ontology Files used in the INFINITECH project.

3.3 Applications

Personal Insurance from Data Analytics (P.I.D.A)

A system to aid customize personalized insurance plans for vehicle owners. Gathering vehicle and trip data from IoT sensors and utilizing a distributed database paired with neural networks, we are able to accurately profile each vehicle owner and provide a suitable insurance plan, including Pay How You Drive (PHYD) and Pay As You Drive (PAYD).

Blockchain-enabled Consent Management

A decentralized and robust blockchain-enabled consent management mechanism, that will enable the sharing of the customers' consent to exchange and utilize their customer data across different banking institutions. It enables the financial institutions to effectively manage and share their customers' consents in a transparent and unambiguous manner. It can store the consents and their complete update history with complete consents' versioning in a secure and trusted manner.

INFINITECH Open API Gateway

The INFINITECH Open API Gateway is a sophisticated API Gateway that encompasses the Open API specification to provide a single point of entry for the added-value functionalities of INFINITECH which are based on microservices.

INFINITECH Data Collection

The INFINITECH Data Collection component is designed and implemented aiming to address the need for a holistic mechanism that will empower the data providers to configure and execute data collection pipelines tailored to their needs.

Scalable Transaction Graph Analysis Component

This application constructs the transaction graph from blockchain transactions and analyses the graph using graph algorithms.

Automatic Data Anonymization Tool for Preserving Privacy and Utility on Datasets

Gradiant's anonymization tool modifies data to preserve privacy. It is especially indicated when a dataset contains personal data, and it has to be outsourced/shared with a third party. It provides different anonymization algorithms that aim at avoiding the appearances of data combinations that could lead to a possible re-identification of the data subjects, while monitoring different privacy and utility metrics to assess the impact of the anonymization process.

3.4 Blockchain

Dynamic Dao Structure for Business Scaling

The dynamic DAO structure aims to solve the following problem: creating an optimal framework for a unicorn DAO. Our solution followed Gall's law, starting from a simple model where within a DAO, small DAOs would be represented by their elected officials to ensure efficiency and quality decision-making.

Gradually, a decentralized monolithic DAO would be achieved using a transparent, interpretable AI-aided decision-making model to determine one's voting quality and power, while optimizing decision quality through accumulated historical data and iterations over time.

BC Based Secure Execution Environment and Data Marketplace for Federated Learning

This work is a collaboration between FBK and IBM within the INFINITECH project. FBK has developed a fraud detection federated learning algorithm, run by

consortium of organizations in their own nodes and on their own data, while sharing the intermediate learning results with each other to achieve higher precision.

IBM provides a blockchain-based secure execution framework for the distributed fraud algorithm execution, recording all the meta-information regarding the execution, and recording the shared intermediate results on a transparent, immutable, and verifiable ledger. At the end of the execution a completed model can be shared as a tradeable asset on a blockchain-based data marketplace also developed by IBM.

ERC1155 Token Smart Contract for Hyperledger Fabric

The Hyperledge Fabric referred as chaincode operates smart contracts in an isolated container environment, such as Docker, and can be written in standard programming languages, such as Go and Node.js. Smart contracts offer the required interfaces that are exploited by applications outside of the blockchain network in order to interact with distributed ledger providing the required level of abstraction, as well as increased level of privacy and confidentiality. To further promote the privacy and confidentiality, Fabric enables the creation of channels in which the participants own a separate ledger of transactions from the rest of the blockchain network that is visible only to the participants of the channel. Finally, it provides the feature of private data, where collections of data can be only be visible and accessible to a portion of the participants of a specific channel. [D4.7]

The implementation and execution of smart contracts varies depending on the blockchain implementation [D4.7]. To facilitate all the operations performed in the ledger within the blockchain network, the smart contracts (or chaincode) are leveraged. Smart contracts are the trusted distributed applications that are deployed within the nodes of the blockchain network and encapsulate the business logic of the blockchain applications. Smart contracts include the agreements that the participants of the blockchain network have formulated with regards to the generation of new facts that are added to the ledger and that will update the current and historical state of the facts that are already stored in the ledger. In this sense, the smart contracts enable the creation of new transactions by the users of the blockchain network by invoking the smart contracts' functions. The smart contracts are facilitating the controlled access to the ledger, offering a layer of abstraction on top of the aspired transactions, encapsulating and simplifying all the relevant information while also ensuring their compliance with the underlying legal agreements, as well as the automation of the several aspects of the transactions.

One of the most critical concepts of blockchain technology is the consensus model that is utilised in order to validate a transaction and to keep the ledger transactions synchronized across the blockchain network. Hence, the consensus model undertakes the validation and approval of the candidate transactions and ensures that the copies of the ledger that are kept within the nodes of the blockchain

network are updated with the same transactions and in the same order. As the blockchain network is composed by multiple nodes, it is very likely that many publishing nodes will compete at the same time to publish new nodes. Additionally, conflicts might be created by nodes publishing new block at approximately the same time. Hence, it is evident that a method is required to ensure that transactions will be written to the ledger at the same order as they generated, as well as that malformed or malicious transactions are rejected. For this reason, the blockchains depending on their implementation specifications exploit different consensus models that are available in computer science such as the CFT (crash fault-tolerant) or BFT (byzantine fault-tolerant) ordering, while at the same time large research effort is spent on this topic towards the definition of further alternative consensus models capable of better addressing this issue with less trade-offs. [D4.7]

3.5 Data Marketplace

BC Based Secure Execution Environment and Data Marketplace for Federated Learning

This work is a joint collaboration between FBK and IBM within the INFINITECH project. FBK has developed a fraud detection federated learning algorithm, run by consortium of organizations in their own nodes and on their own data, while sharing the intermediate learning results with each other to achieve higher precision.

IBM provides a blockchain-based secure execution framework for the distributed fraud algorithm execution, recording all the meta-information regarding the execution, and recording the shared intermediate results on a transparent, immutable, and verifiable ledger. At the end of the execution a completed model can be shared as a tradeable asset on a blockchain-based data marketplace also developed by IBM.

The blockchain-based data marketplace provides a tradeable assets catalogue, where consumer organizations can search for available assets and gain access by paying for those assets using tokens.

3.6 Database Management Systems

Ultra Scalable Transactional Processing Without Blocking in the Advent of Failures

The invented transactional processing method enables to scale horizontally in a linear way transactional management and does not block transactional management when a node fails, thus enabling all transactions to progress except for the ones involved with data of the failed node.

3.7 Dataset

Dataset with Risk Estimates of Major Currency Pairs on the Forex Market

This dataset includes Value at Risk (VaR) and Expected Shortfall (ES) estimations of the major currency pairs on the Forex market. Notably, it provides daily VaR and ES estimates for the AUDUSD, EURCAD, EURCHF, EURUSD, GBPUSD, and USDJPY FX assets for January 2021 to September 2022.

Transaction Graph Dataset for the Bitcoin Blockchain

This dataset contains bitcoin transfer transactions extracted from the Bitcoin Mainnet blockchain.

High Frequency (Tick) Data of Historical FOREX Prices

Price tick data for the most liquid Forex assets (AUDUSD, EURCAD, EURCHF, EURUSD, GBPUSD, USDJPY). The period covered: 09 March 2020 to 07 September 2022.

Synthetic Data of Transactions for Loan Fraud

The data set includes tagged synthetic data of transactions. Tags indicate transactions of fraudulent immediate loan requests.

AEMET Weather Dataset

Dataset composed by AEMET's weather data.

CTAG Alerts Dataset

Traffic alerts data extracted from CTAG stations during the duration of the pilot, describing the category of the alert, location, and timestamp.

SUMO Vigo Vehicles Sample

Dataset compiling a comprehensive set of Vigo vehicle data generated by the SUMO urban mobility simulator.

NBG Datasets – Card Transactions Dataset

Card Transactions Dataset.

NBG Datasets – Card Transactions Dataset

Deposit Account Transactions Dataset.

Pagerank Dataset for Bitcoin Blockchain

This dataset contains Pagerank values and rankings for Bitcoin blockchain addresses and transaction IDs.

Pagerank Dataset for Ethereum Blockchain

This dataset contains Pagerank values and rankings for Ethereum blockchain addresses.

NBG Datasets – Data Dictionary

Data Dictionary for the datasets.

NBG Datasets

Investment Transactions Dataset.

Transaction Graph Dataset for the Ethereum Blockchain

This dataset contains ether as well as popular ERC20 token transfer transactions extracted from the Ethereum Mainnet blockchain.

Processed Synthetic RWD for Tristate Modelling

Processed version of the Raw Synthetic RWD into weekly vectors for training, validation, and testing, including 15 input attributes and a tristate outcome attribute.

Processed Synthetic RWD for Binary Modelling

Processed version of the Raw Synthetic RWD into weekly vectors for training, validation, and testing, including 15 input attributes and a binary outcome attribute.

Raw Synthetic RWD

Real World Data from a simulator for 1,000 people belonging to any of four behavioural groups: athletic, normal, unfit and feeble, simulated over 2 years and 3 months.

3.8 Docker

Data Protection Orchestrator – DPO

The Data Protection Orchestrator orchestrates privacy, security and data protection components and services.

3.8.1 Docker Containers

Pseudoanonymizer

The main purpose of the service is to mask the personal information about the clients and still be informative enough for the purpose of analysis. The service supports different types of conversion for different field types. Text, BIC and IBAN fields: Encryption using a hash or block cypher function.

With this method, the original information is hidden while the same values are preserved after encryption. Numeric and timestamps: rounding and inclusion of noise. Base units for timestamps are days.

The service works by sending an HTTP POST request to [website-address]/pseudonymize with file and data key values. The data is processed as specified in the flow and returned in the same format in which it was received.

Blockchain-enabled Consent Management

A decentralized and robust blockchain-enabled consent management mechanism, that will enable the sharing of the customers' consent to exchange and utilize their customer data across different banking institutions. It enables the financial institutions to effectively manage and share their customers' consents in a transparent and unambiguous manner. It is capable of storing the consents and their complete update history with complete consents' versioning in a secure and trusted manner.

INFINITECH Open API Gateway

The INFINITECH Open API Gateway is a sophisticated API Server that encompasses the Open API specification to provide a single entry point for the added-value functionalities of INFINITECH which are based on microservices.

INFINITECH Data Collection

The INFINITECH Data Collection component is designed and implemented aiming to address the need for a holistic mechanism that will empower the data providers to configure and execute data collection pipelines tailored to their needs.

Decode – Metadata Creation for Banking Domain

Extraction of complex metadata from raw bank documents is paramount to support intelligent data indexing, to face the challenge of sharing info in an effective way within large organisations. Banking language is very specific and rather different from common language. General purpose semantic engines may be not effective in understanding banking related concepts. This evidence raises the need to develop innovative solutions for metadata extraction.

The asset that has been developed is based on a weakly-supervised neural methodology for creating semantic metadata from bank documents. It exploits a neural pre-training method optimized against legacy semantic resources able to minimize the training effort. The method has been tested on documents from the Italian banking community.

3.9 Finance

AI Based Agentive Bank Clerk and Financial Consultant

AI-based agent bank clerk and financial consultant solution provides natural language processing functionalities, allowing users to complete tasks and access information in a more intuitive and user-friendly way.

This could potentially improve market awareness and efficiency by providing faster and more accurate responses to customer inquiries.

FinFlink: Streaming Technical Indicator Generation for Apache Flink

Most financial analytics and financial machine learned models do not process trading data raw, but instead convert that data into a series of more meaningful 'technical indicators' that capture price movements and trends. FinFlink is a Java library which was developed during the INFINITECH project and implements real-time technical indicators within the Apache Flink distributed processing platform.

SME Transaction Categorization

The classification of SME transactions is vital for the additional development of financial management microservices. The absence of labelled data is the main challenge when developing a transaction categorization model.

This model is created by initially hand-labelling a representative subset based on expert knowledge creating a rule-based model, which can then be integrated with a supervised machine learning model, offering a high degree of update automation and transaction re-classification.

PSD2 Open Banking Chatbot SDK powered by botakis

Botakis is an AI powered intelligent agent platform to support automated operations in a conversational way. It is compatible with any device and OS, messaging platforms based on open-source software. The SDK for the PSD2 chatbot is available through the INFINITECH project.

Using the following SDK, there can be an interface with an open banking API of a Bank with the final result of creating a chatbot, through which the user can give consent to execute AIS or PIS scenarios depending on the case.

BC Based Secure Execution Environment and Data Marketplace for Federated Learning

This work is a joint collaboration between FBK and IBM within the INFINITECH project. FBK has developed a fraud detection federated learning algorithm, run by consortium of organizations in their own nodes and on their own data, while sharing the intermediate learning results with each other to achieve higher precision.

IBM provides a blockchain-based secure execution framework for the distributed fraud algorithm execution, recording all the meta-information regarding the execution, and recording the shared intermediate results on a transparent, immutable, and verifiable ledger. At the end of the execution a completed model can be shared as a tradeable asset on a blockchain-based data marketplace also developed by IBM.

3.10 Financial Risk Estimation

DeepVaR: Value-at-Risk Prediction Leveraging Deep Learning

This notebook both explains and implements the DeepVaR, which is a Value-at-Risk model based on deep neural networks and Monte Carlo simulations. The DNN is used to estimate the parameters of the portfolio returns' distribution, which are used to produce the MC samples. As far as the DNN is concerned, the

DeepAR estimator from GluonTS package is utilized to perform probabilistic forecasts, while the VaR is calculated leveraging DeepAR's output.

3.11 Fraud Detection

ML-based Fraud Detection System

ML-based Fraud Detection System goal is to detect frauds in the banking industry related with the request of loan for vehicle purchase. This tool simulates an ML-Based system that has been integrated into an API and connected to a UI. This asset is a result of the participation in the INFINITECH Datathon.

AI Based Detection of Fraudulent Immediate Loan Requests

The asset demonstrates the capabilities of an AI based model detecting fraudulent requests for immediate loans. The AI model facilitates quick analysis of requests, and this way reduces the workload of bank's SoC analysts.

With the help of the model and a lightweight visualization presenting the scoring as well as providing explanations the SoC analyst is enabled to access information quickly and decision in combination with other security tools will be simplified and accelerated.

Fraud Detection System

Robust fraud detection system that aims at improving the detection rate of malicious events (i.e., fraud attempts) and enabling the identification of security-related anomalies while they are occurring by the analysis in real-time of the financial transactions of a home and mobile banking system.

Graph Anomaly Detection

Exploration of Data analytics and advanced machine learning techniques to detect anomalies in bitcoin transaction data. Developed this high-quality code utilizing python and advanced Neural Network libraries such as Pytorch and Pytorch Geometric. The Business model is a pay-as-you-go and it is under development.

3.12 Mobile Application

DUOS – Digital User Onboarding System

Digital User Onboarding System – DUOS Mobile application for dealing with virtual identities in a mobile device.

3.13 Model

Binary Wellbeing Assessment RF Model

Random Forest model to predict short term wellbeing variation, learnt from the "Processed Synthetic RWD for binary modelling" dataset, yielding 76.7% balanced accuracy.

Tristate Wellbeing Assessment RF Model

Random Forest model to predict short term wellbeing variation, learnt from the "Processed Synthetic RWD for tristate modelling" dataset, yielding 64.2% balanced accuracy.

3.14 Monitoring Tool

Blue Behavior

It is a prediction model using satellite data, logbook data, vessel monitoring and staff information, etc. combined to rate vessel – and fisherman behavior. When the model measures a higher probability of illegal activity, a higher probability of inspection by the authorities is triggered. Thus, decreasing the incentive for vessels and fisherman to do illegal fishery.

3.15 Notebook

Graph-based Anomaly Exploration Platform

Graph-based Anomaly Exploration Platform as Jupyter Server engine for data scientists. It is a server-based jupyter notebook which allows expert data professionals to develop scripts and perform a more customized analysis.

The data specialist can interact with algorithm outputs or payments data: it is possible for him/her to write code and therefore explore results with customized queries and/or algorithms.

SMEs Cashflow Prediction

This notebook demonstrates in an explanatory way how to predict in a probabilistic way the future outflows and inflows based on historical transactions. It is applied on SME data, but it can be fine-tuned and applied in general for cashflow prediction.

Portfolio-Value-at-Risk-estimation

This notebook demonstrates in an explanatory way how to estimate and evaluate the Value-at-Risk (VaR) of financial portfolios.

3.16 Recommender

Financial Asset Recommender: Profitability Estimation with Sentiment

This is a supervised financial asset recommender service. It takes as input historical asset pricing data (daily market closing prices), in addition to sentiment indicators derived from news headlines that are mapped to those assets and it predicts the future price for that asset (one month in advance). This is a publicly available containerized service. This component comes in two main flavors, namely:

1. a python jupyter notebook tutorial,
2. a containerized microservice with Rest API that can host a previously trained recommendation model.

Typically, new users should start with the python jupyter notebook tutorial as this will guide you through the process of training a recommendation model based on past data and evaluating its performance. Users with already trained models can then try the microservice container for hosting the models that they have trained.

Financial Asset Recommender: Profitability Estimation

This is a supervised financial asset recommender service. It takes as input historical asset pricing data (weekly market closing prices) and for an asset and it predicts the annualized return on investment (i.e. profit %) for that asset. This is a publicly available containerized service. This component comes in two main flavors, namely:

1. a python jupyter notebook tutorial,
2. a containerized microservice with Rest API that can host a previously trained recommendation model.

Typically, new users should start with the python jupyter notebook tutorial as this will guide you through the process of training a recommendation model based on past data and evaluating its performance. Users with already trained models can then try the microservice container for hosting the models that they have trained.

3.17 Screening Tool

Screening Tool Framework

Anomaly, pattern detection and SCT framework encompass all building blocks for BSI testbed and provide basic functionality for the screening tool. The provided dockerfile imports transaction data, metadata, fuses the data together and prepares a suitable data model and imports it for further use.

3.18 Service

Sentiment Analysis in Financial News

This service performs sentiment analysis in financial news using the FinBERT pre-trained model provided by [ProsusAI and Hugging Face] (https://huggingface.co/ProsusAI/finbert). FinBERT is a pre-trained NLP model to analyze sentiment of financial text. It is built by further training the BERT language model in the finance domain, using a large financial corpus, and thereby fine-tuning it for financial sentiment classification.

3.19 SMEs Business Risks

SMEs Dataset Business Risk

Data set composed of variables that help predict different risks of SMEs.

3.20 Software

AI Profiler Service

The service uses a KMeans clustering method to separate the different driving routes into clusters according to the different values provided by the connected cars. The resulting clusters will divide into each route, allowing the insurance companies to better calculate the prices of the insurance policies. The repository contains the required files to deploy the service providing the driver profiler AI component.

Smart Fleets Platform

Repository hosting the Smart Fleets platform images and necessary deployment files.

3.21 TAHO

The Traffic Analysis Hub Ontology

TAHO is an ontology to represent the Traffic Analysis Hubs data as a semantic model.

3.22 Visualization

Stream Story

Stream Story is a tool intended to help with the analysis and interpretation of time varying data.The system helps the user search for recurrent patterns by representing the data as a diagram of states and transitions, where recurrent patterns visually stand out.

To construct its representation, Stream Story uses and adapts several machine learning techniques. The Stream Story's mechanisms to uncover and explain the structure within the data are visual (hierarchical Markov chain, charts, decision trees, parallel coordinates) and provide a textual narrative explaining/summarising states and patterns within the data.

Veesualive

Business intelligence platform that provides advanced data exploration and visualisation functionalities

3.23 Other

Tokenization on Hyperledger Fabric -ERC20 Chaincode

Tokenization on Hyperledger Fabric -ERC20 chaincode.

This page intentionally left blank

DOI: 10.1561/9781638282334.ch4

Chapter 4

INFINITECH Continuous Integration/ Continuous Deployment Tools

4.1 Tools and Techniques for Testbeds and Sandboxes

This chapter describes the tools and techniques that will be leveraged to implement the testbeds and sandboxes concepts within the INFINITECH project, considering that the INFINITECH-RA is designed leveraging a paradigm based on a microservices [1] architecture implementation, with services interacting among them through REST APIs.

Key pillars for the implementation and deployment of a microservices based architecture are the containers technology [2] and its leading open-source containers orchestration solution, i.e. Kubernetes [3]. Therefore, in order to understand such methodological and technological choices for the realization of the INFINITECH-RA, some key concepts related to containers, microservices and Kubernetes, and how their appearance in the IT environments has deeply changed the software development approaches, are presented.

Finally, the details about how the concrete usage of such technologies benefits and enables the definition of the INFINITECH testbeds and sandboxes concepts are also provided.

4.1.1 Containers Benefits

In the last few years there has been a strong transformation, similar to what happened in the early 2000s with the advent of virtualization, due to containers spread which led to rethinking both how to manage the infrastructure and how to design and build the applications.

The advent and the diffusion of containers has made it possible to improve the computational management of infrastructures, thanks to the possibility of removing the overhead generated using the hypervisor (integrated software that allows to virtualize the HW resources of a server and make them available among several applications), and through the usage of the functionalities already available within the Linux OS kernel, as in Figure 4.1.

The container engine, to date the most used one is Docker [4], takes advantage of

- Linux-Control groups: Allow each container to get its fair share of memory, CPU, disk I/O and Network stack. At the same time a single container cannot bring the system down by exhausting one of those resources.
- Linux-Namespaces: Provide the possibility to isolate the processes running into a container from processes running in another container or in the host system.

to run containers like VMs isolated each other.

In the end, a container is a package that contains code, system libraries, dependencies and software tools. An example is given in Figure 4.2.

The main advantages to use containers respect VMs can be summarized in these macro points:

- Size: a Container is small.
- Overhead: no fully OS is required.
- Speed: Boot time is faster.
- Scaling: Real time provisioning.

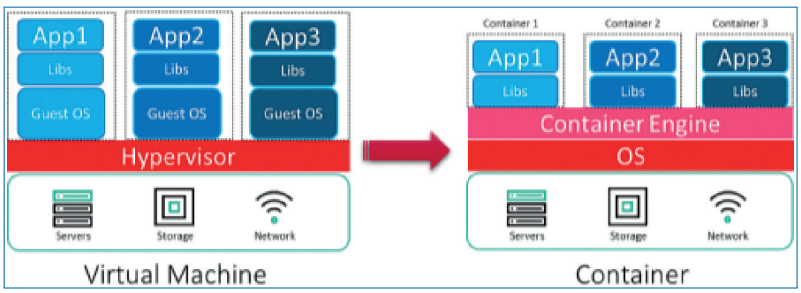


Figure 4.1. VM vs container.

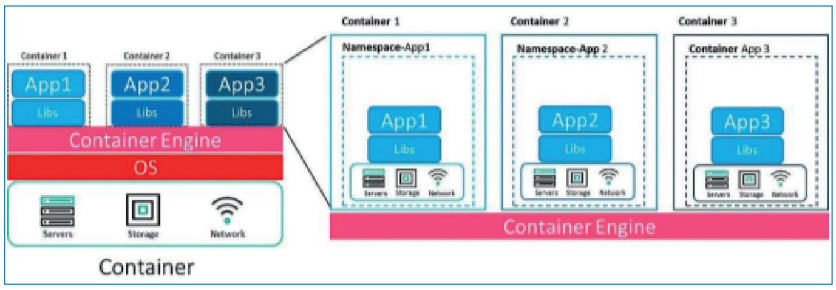


Figure 4.2. Container kernel properties.

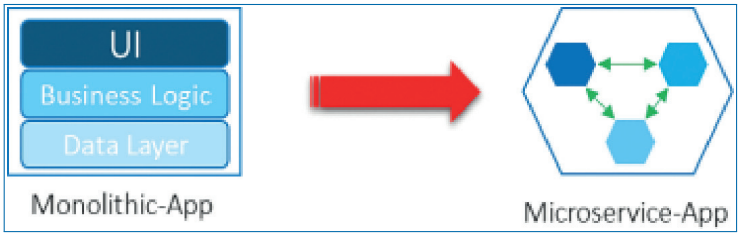


Figure 4.3. Monolithic vs microservice.

At the same time, in order to take advantage of containers is necessary to rethink and redesign the currently monolithic applications into microservices based applications.

4.1.2 Microservices Approach

The spread of the containers led, as already mentioned, to the necessity to change the approach of the developers in the creation of an application, moving from design of monolithic applications, where the various components (generally UI, Business logic and Data-layer) were strongly coupled among them, to microservices applications where the various components (i.e. microservices) are decoupled from each other (see Figure 4.3).

This methodological change has both advantages and disadvantages.

The advantages can be summarized as:

- Simple to develop: Each microservice is independent and small.
- Simple to upgrade: Since each microservice is independent it's possible to upgrade each component independently.
- Simple interaction: Each microservice communicates with each other through well-defined and standard interfaces (API).
- Simple to scale: Each microservice can be scaled independently.
- Scale on demand: It's possible to run multiple microservices behind a load balancer to scale on demand request.

The disadvantages can be summarized as:

- Complexity: Splitting an application into multiple independent microservices increases the complexity of the deployment process.
- Monitoring: Monitoring each microservice requires to have many metrics and logs to manage it.
- Performance: All communications occur on network so they are slower than memory communications.
- Debugging: a Monolithic application is much easier to debug and test due to the fact it is composed by single indivisible units.

Our approach for a rapid iterative development of a microservice based software infrastructure builds upon a widely used methodology like DevOps.

4.1.3 Kubernetes Containers Orchestration

The arrival on the market of containers and related microservices based applications, on the one hand enabled applications that quickly scale according to the requirements and that could be easily updated, on the other hand meant that software previously managed as a single indivisible piece was split into several dozens of microservices (containers), making it more difficult to manage them.

In this context, the necessity to develop a tool that was able to manage the lifecycle of the microservices (deployment, scaling, and management) arose: such tool was developed by Google with the name of "Project Seven of Nine" and released as open source software in 2014. Today such tool is widely known as Kubernetes.

Kubernetes provides:

- Service discovery and load balancing. Kubernetes can expose a container using the DNS name or using its own IP address. If traffic to a container is high, Kubernetes can load balance and distribute the network traffic so that the deployment is stable.
- Storage orchestration. Kubernetes allows to automatically mount a storage system of different types, such as local storages, public cloud providers, and more.
- Automated rollouts and rollbacks. You can describe the desired state for your deployed containers using Kubernetes, and it can change the actual state to the desired state at a controlled rate. For example, you can automate Kubernetes to create new containers for your deployment, remove existing containers and adopt all their resources to the new containers.
- Automatic bin packing. You provide Kubernetes with a cluster of nodes that it can use to run containerized tasks. You tell Kubernetes how much CPU and memory (RAM) each container needs. Kubernetes can fit containers onto your nodes to make the best use of your resources.

- Self-healing. Kubernetes restarts containers that fail, replaces containers, kills containers that do not respond to your user-defined health check, and does not advertise them to clients until they are ready to serve.
- Secret and configuration management. Kubernetes lets you to store and manage sensitive information, such as passwords, OAuth tokens, and SSH keys. You can deploy and update secrets and application configuration without rebuilding your container images, and without exposing secrets in your stack configuration.

4.1.4 Kubernetes Architecture

A Kubernetes (aka K8s) cluster is made up of two macro blocks, the first one called Control Plane (Master) and the second one called Data Plane (Worker).

The Control Plane constitutes the brain of the cluster and internally it is made up of the following components:

- Kube-APIserver is the component that exposes cluster API and truly is the main component, since Kubernetes has been designed and built to base all the operations on the use of the API.
- Etcd is a key value database that maintains all information relating to the status of the cluster.
- Kube-scheduler schedules on which nodes of the Data Plane runs the containers (in Kubernetes named POD, the smallest deployable units of computing that can be created and managed in Kubernetes) based on the resources required, cluster status and affinity and anti-affinity rules.
- Kube-controllermanager consists of a set of control processes which:
  - Check if cluster nodes are active.
  - Check if number and status of running POD it's required one.
  - Control and create token to access on the K8s resource.
  - Populates the Endpoints object (that is, joins Services & Pods).

The Data Plane is the part where the workload is carried out, i.e. where the PODs are put into execution and it is characterized by the following components:

- Kube-proxy is a proxy that allows communication to the PODs from within and outside the cluster.
- Kubelet checks that PODs are running.
- Container runtime (engine) is a software responsible for running containers in Kubernetes can be used(Docker, CRI-O and containerd).

In the Figure 4.4 the Kubernetes Cluster components is depicted:

For an outline of the entire Kubernetes solution, see the documentation on kubernetes.io.

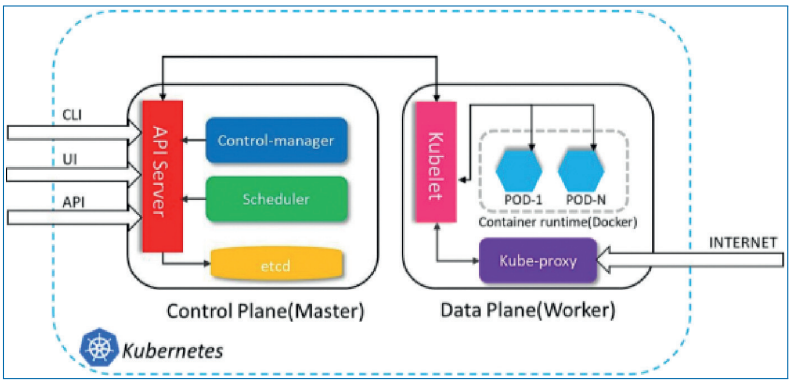


Figure 4.4. Kubernetes architecture.

However, for the purpose of this chapter it is sufficient to now outline two main concepts available in Kubernetes that we will use later to implement the Sandbox concept.

1. Namespaces: They are a logical grouping of a set of Kubernetes objects to whom it's possible to apply some policies, in particular:
   - Quote sets the limits on how many HW resources can be consumed by all objects.
   - Network defines if the namespace can be accessed or can access to other Namespaces, in other word if the Namespace is isolated or accessible. Different policies can be given to different namespaces.
2. POD (Figure 4.5) is the simplest unit in the Kubernetes object. A Pod encapsulates one container, but in some cases (when the application is complex)

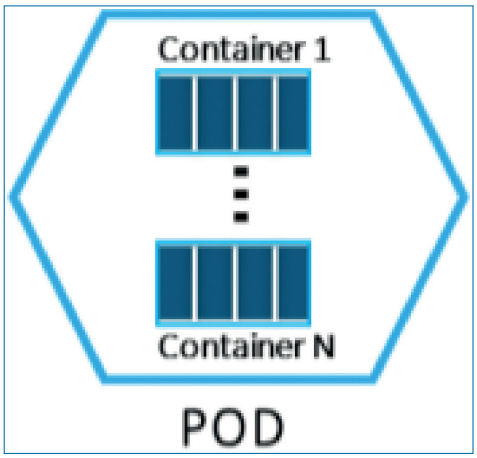


Figure 4.5. Kubernetes POD.

a POD can encapsulate more than one container. Each POD has its own storage resources, a unique network IP, access port and options related to how the container/s should run.

4.2 INFINITECH Testbeds

INFINITECH will make available a number of testbeds for experimentation, testing and validation of BigData, AI and IoT solutions, including:

- 10 testbeds that are established in the data centres of incumbent financial organizations(on premise testbeds).
- 1 testbed that will be provisioned and established in order to support the experimentation of the FinTech and InsurTech pilots and enterprises of the consortium, hosted on the partner NOVA's Data Center.
- 1 testbed that will be provisioned and established in order to support the experimentation of the INFINITECH blueprint reference testbed (see Section 4.2 of the deliverable, associated to one of the official INFINITECH pilots), hosted on the AWS (Amazon Web Services) [6] public provider.

Accordingly, the current, (at the time of writing, subject to possible evolutions along the project lifecycle), INFINITECH project plan is to deliver 15 pilots: 10 out of 15 will be carried out on dedicated on premise Data Centres, while the remaining 5 out of 15 will be carried out on the NOVA's Data Centre, a shared INFINITECH Data Centre. In addition, a blueprint reference testbed will be also provided, built upon the requirements of one of the INFINITECH pilots, as stated above.

The set of hardware resources like storage, compute and network will be considered a testbed, as shown in Figure 4.6.

It is not relevant where these resources are deployed: they can be inside a private Data Centre or in any cloud provider.

Therefore, the 15 pilots that have been foreseen will be executed in 10+2 testbeds, in addition to the blueprint reference testbed, as shown in Figure 4.7.

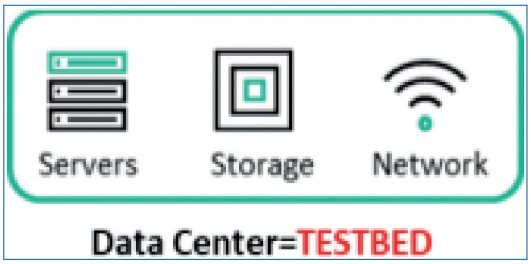


Figure 4.6. Testbed.

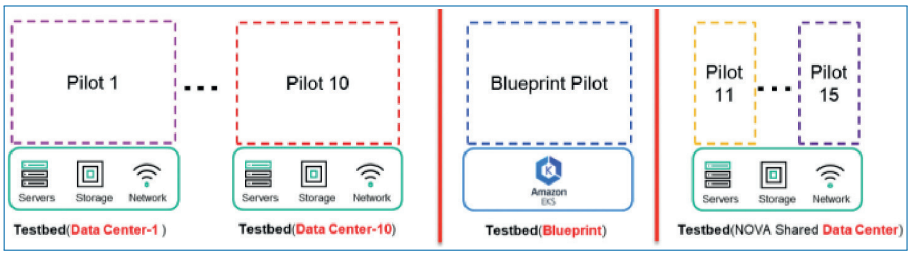


Figure 4.7. Testbeds and pilots.

4.3 INFINITECH Sandboxes

Each INFINITECH pilot will have one or more Use Cases (realized by one or more pilot Apps, each one realized by one or more INFINITECH microservices): in our vision each Use Case will be a Sandbox provisioned by the leverage of Kubernetes Namespaces.

In fact, as we already said in the previous paragraph, the Kubernetes Namespace feature makes it possible to logically isolate the objects (mainly PODs) inside it from other Namespaces. Therefore, each dedicated Testbed will only have one Kubernetes cluster with as many Namespaces as the number of Use Cases to be implemented for a single pilot (see Figure 4.8). In the other case, each (2 out 10)

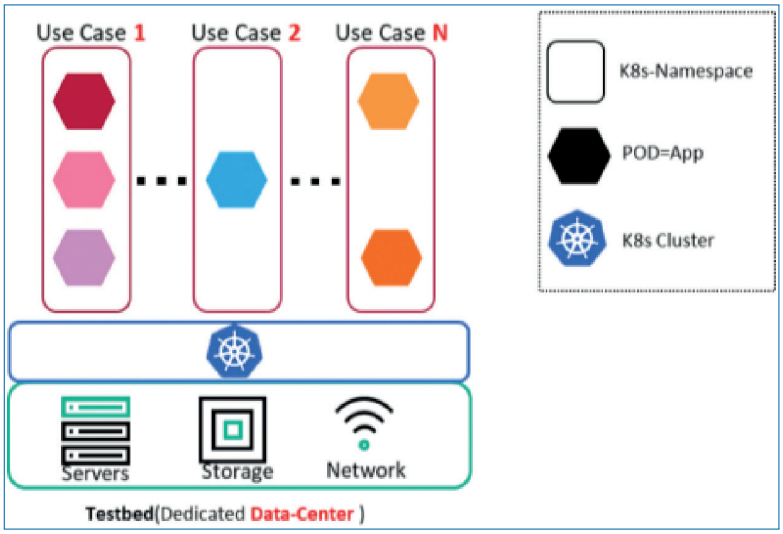


Figure 4.8. Sandboxes in a dedicated Testbed.

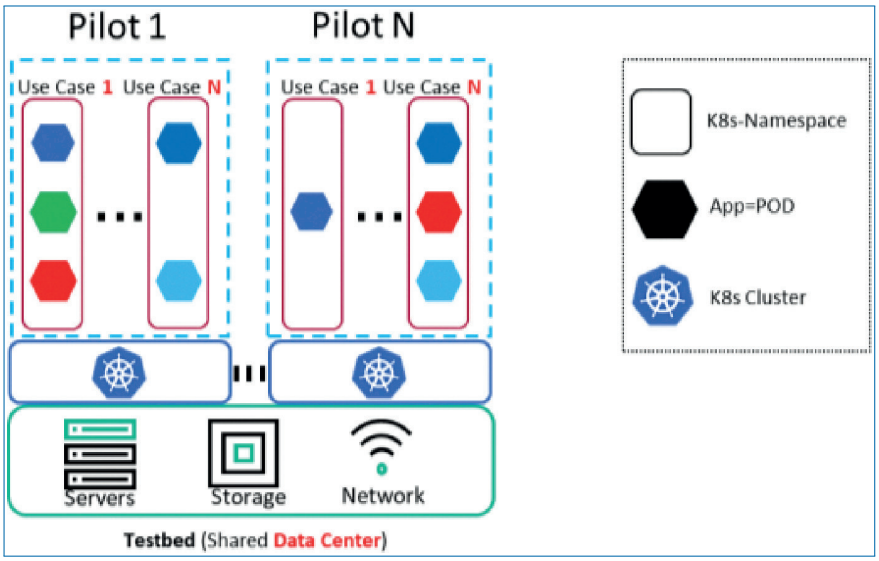


Figure 4.9. Sandboxes in a shared Testbed.

shared Testbed will have one Kubernetes cluster for each pilot it has to host and manage (see Figure 4.9).

In the end, in this general context, each pilot App will be realized as a Kubernetes POD.

This page intentionally left blank

DOI: 10.1561/9781638282334.ch5

Chapter 5

INFINITECH Tools and Techniques for Management of Data Sets

5.1 Tools and Techniques for Management of Datasets

This chapter describes the tools and techniques that will be leveraged for the management of datasets within the INFINITECH project and how they are linked and mapped with the concepts and techniques related to testbeds and sandboxes and also with respect to the blueprint reference testbed environment.

5.1.1 Data Sources and Data Access

The INFINITECH platform aims to host and serve the needs of a variety of applications and tools used by the finance and insurance sectors, which have diverse needs and requirements for data access. We envisage that different types of data sources will be supported such as:

- structural, semi-structural or completely un-structured data,
- static data, often called data at-rest or streaming data, while the source might be stored on premises, in a third-party organization or should be imported inside the deployed sandbox. For instance, a common solution might need

to use historical data in order to perform post processing and apply ML/DL analytics for risk assessment.

However, modern enterprises are not only interested in extracting knowledge and identifying risks or opportunities based on historical and often obsolete data, but they also need to extract this type of information from real data, which introduces a significant challenge regarding data management from the tailored sandbox perspective. To further complicate the needs of the platform, other requirements coming from the finance and insurance sectors are the identification of potential fraudulent financial transactions as they occur, or pre-processing of data from IoT devices as it arrives. Data analysts often tend to use different datasets that can be retrieved from third party organizations or other external sources such as social media to enrich their algorithms with deeper knowledge in order to build a more accurate profile of their customers, either referring to individuals or to enterprises. Finally, additional considerations must be taken into account when dealing with the volume and privacy of the data of a finance organization, where moving data from the source to a sandbox might not be feasible either due to regulatory constraints, or simply due to the overall volume of the datasets.

In order to address all these requirements, the INFINITECH-RA has been designed in order to take into account the different types of data sources. With respect to the implementation to support the functionality needed for data management and processing, details have been given in the corresponding deliverables of WP3, WP4 and WP5.

In looking at the techniques and tools that will allow the tailored sandboxes to enable the data management of these diverse environments, the focus is on data access. In other words, how the different components that will be deployed inside the sandbox can be integrated in a way that allows them to make use of all the aforementioned data sources.

We can identify 6 different focuses for data access:

- Static data ingestion.
- Dynamic data ingestion.
- IoT streaming.
- Direct access to on premise data sources.
- Blockchain data access.
- Third party data access.

In the following subsections, we will describe the details of each of those different means of data access and will analyze the requirements and technical considerations that need to be tackled by the infrastructure, when orchestrating the deployment and maintenance of the integrated solution. The next section will give more insights

on how the INFINITECH platform overcomes the barriers and deals with the challenge of securely allowing those diverse means for data access, thus enabling the automated deployment of the integrated solution.

5.1.2 Static Data Ingestion

This is the most typical case for data access that is foreseen to be a requirement for the majority of the cases that will need to be deployed into an INFINITECH sandbox. The data analyst has their own data stored into a database management system or other form of persistent storage and needs to extract a specific dataset and ingest it into the central data repository deployed inside the sandbox so that they can use the tools provided by the platform to perform the analysis. They will have to extract this information in a static file in a specific format (i.e. csv) and trigger the data ingestion process of INFINITECH. We call this data access mode static data ingestion, as the file will be created once, and will be used in each deployment. This access mode also covers the needs for a data analyst to make use of an existing synthetic dataset that has been provided and made available to the platform, via the INFINITECH marketplace. The generated file must be stored into a persistent storage volume and the latter must be visible from the sandbox so that the tools deployed inside can have access to it. This mainly refers to the data ingestion component that will take care of the data migration process, possibly applying a pre-processing algorithm to clean, harmonize and anonymize the dataset and finally store the raw data into the data repository. It will also support cases where direct access to the static file might be needed by the analytical tools on the analytics layer of the RA, where there is no need for additional processing by the lower logical levels of the architecture (i.e. a spark job needs to grab everything from a csv file, where no additional processing like a filter or join operation can be pushed down to the processing layers).

5.1.3 Dynamic Data Ingestion

In this scenario, data ingestion does not take place once in the initialization of the sandbox, rather the data might be migrated periodically to the data repository. We call this dynamic data ingestion as the dataset inside the sandbox will be updated with new data after a given period of time. This will cover cases where the requirement is to have the sandbox deployed, integrated and synchronized with the data sources of the organization. For instance, the integrated solution might need to get updated with the current snapshot of the data, so that the data analyst can benefit from a daily picture of the transactions of any given customer. Rather than having a human intervene to manually extract the data needed into

a static file that will be loaded into the system, the data provider can implement specific APIs so that the data migration can be done in a fully automated manner every time (i.e. end of the business day). The data ingestion component of INFINITECH will need to open a connection to those APIs in order to retrieve the data according to the given specified protocol or configuration. The APIs can vary, from REST implementations that imply HTTP connections, to database specific ones (JDBC, ODBC, etc.) that will imply TPC connections, or even SFTP connections to a file server. This introduces the requirement for the sandbox to allow the INFINITECH component responsible for this job to access endpoints outside the sandbox using different communication protocols. Regarding the need for data access, only the data ingestion component will need to have access to those endpoints, in order to migrate the data into the central data repository. All other components in the different logical layers of the RA will retrieve the data via the data repository.

5.1.4 IoT Streaming

As we already mentioned, it becomes more crucial for modern applications in the finance and insurance sector to perform real-time analysis in order to respond to possible risks, identify opportunities or even detect fraud transactions, as they occur. This type differentiates from the others as it does not rely on data at-rest, but on streaming. As a result, content event processing over a data stream is becoming popular. Data transmitted over a stream is usually small and contains information from either a sensor deployed in a vehicle or in the soil, information coming from a finance transaction, logging information produced when a user is navigating the web or even a tweet or post on social media. We consider all these types of data as IoT, and we categorize this type for data access as IoT Streaming. The unified query processing framework of INFINITECH will require access to data streams, as it provides the means to deploy continuous queries that can perform live processing over the content of the streams, possibly making use of the other layers that allow the combination of operators targeting live data with static data at-rest. The unified query processing framework contains a streaming engine that consumes the stream from the source. The requirement for the infrastructure in this case is to allow the components that are deployed inside the sandbox to be accessible from outside of the sandbox. To concretize the scenario, the sources of the data streams must connect to the query processing framework of the platform by establishing static TPC connections to the latter. All other components in the different logical layers of the RA will retrieve data and information from this layer, so there is no need for others to be accessible from external data sources.

5.1.5 Direct Access to on Premise Data Sources

This covers scenarios where it is not enough for the integrated solution to have access only to a snapshot of the dataset that has been periodically loaded into the sandbox, but it requires access to the overall dataset. It will also cover scenarios where data cannot be migrated into the sandbox due to their volume or due to other regulatory constraints. In the case of 'big data' management, it might be better to push down the data processing to the source and grab only the results that will feed the tools and components in the analytical layer, rather than moving all the data inside the sandbox to be loaded to the central repository beforehand. We call this mode for data access as direct access to on premise sources as the sandbox needs to access directly the data source and send the processing there, and it will not migrate any data inside. The polyglot component of INFINITECH is responsible for this functionality, which lies in the processing layer of the RA, even if it pushes down to processing to the source. However, when the processing takes place, from a logical point of view, the polyglot takes care of this and all other layers in the analytical layers will make use of the latter. The requirements for the sandbox are similar with the direct data ingestion mode, but it is the polyglot component now that needs to open connections to external endpoints from the sandbox, using different communication protocols, varying from HTTP, to TPC and SFTP.

5.1.6 Blockchain Data Access

As the name of this access mode implies, it covers scenarios where secure access to data stored in the Blockchain is required. We are interested here in uses of Blockchain as a means of persistent storage of data, and not as a means of verifying the consensus for a given transaction. As it has been described in the corresponding deliverables of the tasks related with the Blockchain technology in WP4, the access is being granted via a specified API. Therefore, from a functional view, the components that need to access Blockchain data will have to open a connection to the provided APIs and retrieve data. The Blockchain data needs no pre-processing to the source neither can it benefit from the processing capabilities of the INFINITECH data management components. Therefore, it is the components that are lying in the analytical layer of the RA that need to be granted access to the Blockchain storage via the appropriate designated endpoint. However, from the perspective of the infrastructure that will orchestrate the automated deployment and maintenance of the integrated solution, the requirements are similar to the dynamic data ingestion: to allow components inside the sandbox to establish connections with endpoints that are located externally from the sandbox.

5.1.7 Third Party Data Access

In this access mode, the components and analytical tools provided by the INFINITECH platform need to retrieve data and information that has been provided by external sources. This scenario is different from the previous ones, and especially to the direct access to on premise sources, as the third party does not grant access to its entire dataset, rather than provides specific APIs that will allow others to extract only specific information. In other words, it can allow to submit a consult (i.e. how popular is the given trend in the articles of UK during the last week) and retrieve its result. The third-party organization has access to other datasets, and has already performed an initial pre-processing that generates the results of the types of consults its API allows to perform. Therefore, there is no need for the data management component that are suited in the processing layers of the RA to perform any further analysis and this information can be directly used by the analytical tools. However, from the perspective of the infrastructure that manages the sandboxes, the requirement is the same as the Blockchain: to allow the components deployed in the sandbox to open connections to APIs that are deployed outside.

5.1.8 Datasets Management from the Blueprint Reference Environment Perspective

Even if we have several and diverse types for data access, the requirements for the datasets and data access management tools and the techniques from the blueprint reference environment perspective, or rather from the testbeds and sandboxes perspective, can be described and addressed by grouping the methodologies presented in the previous paragraphs into the following macro categories:

1. Static data ingestion: The blueprint must be able to store static data into a persistent volume that needs to be accessible by the sandboxes. This requirement is fulfilled using the Kubernetes PV (Persistent Volume) available inside the blueprint and in particular by using the PVC (Persistent Volume Claim) that allows to split the PV according to the sandboxes needs.
2. Dynamic data ingestion – Blockchain data access – Third party data access: The blueprint must be able to allow the components of different layers of the logical architecture view to open connections to external endpoints. Accordingly, the components deployed inside the sandboxes need to communicate with the external endpoints. This communication is permitted by default, unless otherwise stated, through the API Gateway which forwards HTTP/HTTPS requests to and from the sandboxes.
3. IoT Streaming – Direct access to on premise data sources: The blueprint must be able to allow for components that have been already deployed to connect

and maintain open connections with components that have been deployed inside a sandbox. Accordingly, the components that have been deployed inside a sandbox have to maintain an open connection on TPC protocol. In this case the connection are not managed by the API Gateway, but using the Kubernetes NodePort. The NodePort functionality exposes the sandbox service on each worker's IP at a static port in the range between 30000 and 32767. Each worker proxies such port (the same port number on every Node) into the sandbox service. In order to clarify this concept, let's suppose that the sandbox service works on the port 390: this port is then remapped on the port 30390 on each worker node, and in this way the sandbox can then be reached by external components and maintains an open connection on the port 30390.

This page intentionally left blank

DOI: 10.1561/9781638282334.ch6

Chapter 6

INFINITECH Reference Testbed and Sandbox

6.1 Overview of the INFINITECH Blueprint Reference Testbed

This chapter describes the initial design and implementation of the INFINITECH blueprint reference testbed, through the actual realization (with a full compliance) of the INFINITECH-RA Development and Deployment views, in terms of the concrete specification and realization of the fundamental and target INFINITECH concepts of Testbeds, Sandboxes and Datasets management, and related tools and techniques for their effective setup and deployment in the INFINITECH pilots and validation scenarios. In other words, how the concept described earlier will be realized on the target INFINITECH infrastructure environments.

Moreover, the planned blueprint environment associated to the initial and preliminary Proof of Concept (PoC) implementation of one of the official INFINITECH pilots (at the time of writing, the WP7 Pilot 5b: Business Financial Management (BFM) tools delivering Smart Business Advice, owned by the partner Bank of Cyprus (BOC)) is also described.

6.1.1 Development View

With respect to the Development view, it has been decided to implement DevOps (Development and Operations) processes. DevOps represents a change in IT culture, focusing on rapid IT service delivery through the adoption of agile, lean practices in the context of a system-oriented approach. DevOps emphasizes people (and culture), and seeks to improve collaboration between operations and development teams. DevOps implementations utilize technology–especially automation tools that can leverage an increasingly programmable and dynamic infrastructure from a life cycle perspective [5].

The practical implementation of DevOps goes through the CI/CD processes, which are delivered through the combined practices of Continuous Integration (CI) and Continuous Delivery (CD). In particular:

- Continuous Integration is a practice where development teams frequently commit (many times per day) application code changes to a shared repository. These changes automatically trigger new builds that are then validated by automated testing (as in Development Testing, DevTest [8]), to ensure that they do not break any functionality.
- Continuous Delivery is an extension of the CI process. It's the automation of the release process so that new code is deployed to target environments, typically to test environments, in a repeatable and automated fashion.

In order to fulfil the INFINITECH project goals, the CI/CD processes have been created in the context of the blueprint reference testbed environment. Moreover, to build a system working consistently as a whole, the developers writing the individual components of the INFINITECH platform need an integrated environment where they can test their components working together with the other services. To support this process, we implemented a Continuous Integration environment based on EKS (Elastic Kubernetes Service) [7], a managed Kubernetes service on the AWS public cloud. More details on EKS will be provided in the following section Kubernetes is an ideal choice for a Continuous Integration environment, since it allows easy updates of deployments when new application images are built, with manifests containing deployment configurations versioned like Git [8] alongside the application source code. Furthermore, it is easy to spin up new test environments from scratch, which enables future scenarios including automated end-to-end integration testing. Build agents are also created on demand and removed when done, providing efficient resource utilization and clean environments to ensure build reproducibility.

On the target cluster, a namespace named devops have been created for hosting the DevOps tools, which are:

- Gitlab [9] is a Git repository manager that lets each developer teams collaborate on the INFINITECH's source code.
- Jenkins [10] is the de-facto standard open source automation server for orchestrating CI/CD workflows.
- Sonatype Nexus [11] is a popular artefact repository that also works as a Docker registry, as required in our case.
- OpenLDAP [12] is used as the single user directory for all tools, centralizing authentication and simplifying management of developer accounts.
- Helm [13] is a package manager that streamlines installing and managing Kubernetes applications.

Figure 6.1 shows how CI/CD works for a specific partner (e.g. Partner "A"). When a developer pushes new component code, Gitlab invokes a webhook on Jenkins, which starts any job affected by the code changes. The job builds the component, runs unit tests and, if everything has worked in a proper way, builds an updated Docker image, and pushes it to Nexus. The following step is deploying the updated component in the specific partner namespace; in fact, we will have as many namespaces as the partners in order to maintain the correct isolation between all INFINITECH partners. In order to deploy the component Helm manager will be used. At the end of the process, Jenkins sends a notification to a dedicated CI/CD channel on the INFINITECH Slack [14] project, so that developers are informed that a new build occurred and whether it was successful or not. In case of errors, developers will have to inspect the build logs, find the problem, and correct it. In case of success, developers will go ahead and test that the new version works correctly in the test environment.

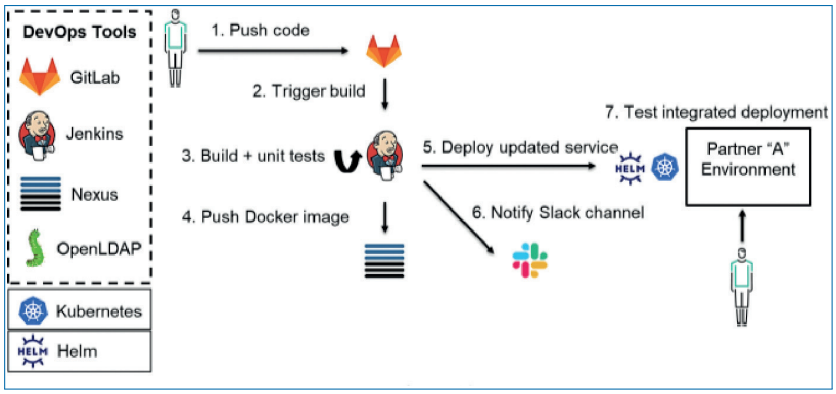


Figure 6.1. CI/CD workflow.

We intend to enhance the process by adopting a DevSecOps [15] approach and including the related tools like Sonarqube [16] in the CI/CD pipeline.

DevSecOps aims to include security in the software development life cycle from the beginning, following the same principles of DevOps. Security is then considered throughout the process and not just as an afterthought at the end of it, so that different kinds of security checks are executed continuously and automatically, giving developers quick feedback if the latest changes introduced a vulnerability that must be corrected.

DevSecOps of course requires that security experts work side by side with developers and operations to make sure that security requirements are addressed and best practices followed, in addition to validating product design and architecture.

Moreover, in this context, in order to facilitate the Machine Learning (ML) development and testing, we have also planned to introduce (again, in the following iterations of WP6 T6.2 and T6.3 tasks and activities), the MLOps [17] (a compound of Machine Learning and IT Operations) methodology focused on:

- Facilitate communication and collaboration between teams.
- Improve model tracking, versioning, monitoring and management.
- Standardize the machine learning process to prepare for increasing regulation and policy.

This add-on enhanced methodology will enable us to automate the porting of the machine learning algorithms, as much as possible, in production environments.

Putting this into practice is often very complicated, because ML processes are often based on heterogeneous environments.

Therefore, the first step towards MLOps requires the standardization of these environments as much as possible, and in this respect the Kubernetes technology and containers provide the abstraction, scalability, portability, and reproducibility required to run the same piece of software in all these environments. As a second step, it is necessary to make standard the workflows used for the construction and building of the ML models.

In this sense, we have selected the Kubeflow [18] platform as candidate for integration in our blueprint reference testbed, in order to provide an infrastructure to build models capable of enabling the portability of these models and workflows. In particular, ML workflows are defined as Kubeflow pipelines and a pipeline consists of these steps:

- Data preparation.
- Training.
- Testing.
- Serving.

Each step is a container, and the output of each step is the input of the following step. Once compiled, this pipeline is portable across different environments.

6.1.2 Creation of the EKS INFINITECH Cluster

The creation of EKS INFINITECH cluster involves four principal steps:

- Create a specific IAM (Identity and Access Management) Role able to create a provisioning EKS cluster.
- Create the VPC (Virtual Private Cloud).
- Create the EKS Control plane.
- Create the Worker node.

To perform all the steps we will leverage the AWS console and where possible also the CloudFormation templates that allow us to replicate the installation and configuration anytime in very straightforward way. For the creation of the IAM role we will use the following Amazon CloudFormation template:

```
---
AWSTemplateFormatVersion: '2010-09-09'
Description: 'Amazon EKS Cluster Role'


Resources:

  eksClusterRole:
    Type: AWS::IAM::Role
    Properties:
      AssumeRolePolicyDocument:
        Version: '2012-10-17'
        Statement:
        - Effect: Allow
          Principal:
            Service:
            - eks.amazonaws.com
          Action:
          - sts:AssumeRole
      ManagedPolicyArns:
        - arn:aws:iam::aws:policy/AmazonEKSClusterPolicy

Outputs:

  RoleArn:
    Description: The role that Amazon EKS will use to create AWS resources for
Kubernetes clusters
    Value: !GetAtt eksClusterRole.Arn
    Export:
      Name: !Sub "${AWS::StackName}-RoleArn"
```

To create the VPC for EKS we will use the following template:

```
---
AWSTemplateFormatVersion: '2010-09-09'
Description: 'Amazon EKS Sample VPC - Private and Public subnets'

Parameters:

  VpcBlock:
    Type: String
    Default: 192.168.0.0/16
    Description: The CIDR range for the VPC. This should be a valid private (RFC
1918) CIDR range.

  PublicSubnet01Block:
    Type: String
    Default: 192.168.0.0/18
    Description: CidrBlock for public subnet 01 within the VPC

  PublicSubnet02Block:
    Type: String
    Default: 192.168.64.0/18
    Description: CidrBlock for public subnet 02 within the VPC

  PrivateSubnet01Block:
    Type: String
    Default: 192.168.128.0/18
    Description: CidrBlock for private subnet 01 within the VPC

  PrivateSubnet02Block:
    Type: String
    Default: 192.168.192.0/18
    Description: CidrBlock for private subnet 02 within the VPC

Metadata:
  AWS::CloudFormation::Interface:
    ParameterGroups:
      -
        Label:
          default: "Worker Network Configuration"
        Parameters:
          - VpcBlock
          - PublicSubnet01Block
          - PublicSubnet02Block
          - PrivateSubnet01Block
          - PrivateSubnet02Block

Resources:
  VPC:
    Type: AWS::EC2::VPC
    Properties:
      CidrBlock:  !Ref VpcBlock
      EnableDnsSupport: true
      EnableDnsHostnames: true
      Tags:
      - Key: Name
        Value: !Sub '${AWS::StackName}-VPC'

  InternetGateway:
    Type: "AWS::EC2::InternetGateway"

  VPCGatewayAttachment:
    Type: "AWS::EC2::VPCGatewayAttachment"
    Properties:
      InternetGatewayId: !Ref InternetGateway
      VpcId: !Ref VPC
```

```
PublicRouteTable:
  Type: AWS::EC2::RouteTable
  Properties:
    VpcId: !Ref VPC
    Tags:
    - Key: Name
      Value: Public Subnets
    - Key: Network
      Value: Public

PrivateRouteTable01:
  Type: AWS::EC2::RouteTable
  Properties:
    VpcId: !Ref VPC
    Tags:
    - Key: Name
      Value: Private Subnet AZ1
    - Key: Network
      Value: Private01

PrivateRouteTable02:
  Type: AWS::EC2::RouteTable
  Properties:
```

```
    VpcId: !Ref VPC
    Tags:
    - Key: Name
      Value: Private Subnet AZ2
    - Key: Network
      Value: Private02

PublicRoute:
  DependsOn: VPCGatewayAttachment
  Type: AWS::EC2::Route
  Properties:
    RouteTableId: !Ref PublicRouteTable
    DestinationCidrBlock: 0.0.0.0/0
    GatewayId: !Ref InternetGateway

PrivateRoute01:
  DependsOn:
  - VPCGatewayAttachment
  - NatGateway01
  Type: AWS::EC2::Route
  Properties:
    RouteTableId: !Ref PrivateRouteTable01
    DestinationCidrBlock: 0.0.0.0/0
    NatGatewayId: !Ref NatGateway01
```

```
PrivateRoute02:
  DependsOn:
  - VPCGatewayAttachment
  - NatGateway02
  Type: AWS::EC2::Route
  Properties:
    RouteTableId: !Ref PrivateRouteTable02
    DestinationCidrBlock: 0.0.0.0/0
    NatGatewayId: !Ref NatGateway02

NatGateway01:
  DependsOn:
  - NatGatewayEIP1
  - PublicSubnet01
  - VPCGatewayAttachment
  Type: AWS::EC2::NatGateway
  Properties:
    AllocationId: !GetAtt 'NatGatewayEIP1.AllocationId'
    SubnetId: !Ref PublicSubnet01
    Tags:
    - Key: Name
      Value: !Sub '${AWS::StackName}-NatGatewayAZ1'

NatGateway02:
  DependsOn:
  - NatGatewayEIP2
  - PublicSubnet02
  - VPCGatewayAttachment
```

```
  Type: AWS::EC2::NatGateway
  Properties:
    AllocationId: !GetAtt 'NatGatewayEIP2.AllocationId'
    SubnetId: !Ref PublicSubnet02
    Tags:
    - Key: Name
      Value: !Sub '${AWS::StackName}-NatGatewayAZ2'

NatGatewayEIP1:
  DependsOn:
  - VPCGatewayAttachment
  Type: 'AWS::EC2::EIP'
  Properties:
    Domain: vpc

NatGatewayEIP2:
  DependsOn:
  - VPCGatewayAttachment
  Type: 'AWS::EC2::EIP'
  Properties:
    Domain: vpc
```

```
PublicSubnet01:
  Type: AWS::EC2::Subnet
  Metadata:
    Comment: Subnet 01
  Properties:
    MapPublicIpOnLaunch: true
    AvailabilityZone:
      Fn::Select:
      - '0'
      - Fn::GetAZs:
          Ref: AWS::Region
    CidrBlock:
      Ref: PublicSubnet01Block
    VpcId:
      Ref: VPC
    Tags:
    - Key: Name
      Value: !Sub "${AWS::StackName}-PublicSubnet01"
    - Key: kubernetes.io/role/elb
      Value: 1

PublicSubnet02:
  Type: AWS::EC2::Subnet
  Metadata:
    Comment: Subnet 02
  Properties:
    MapPublicIpOnLaunch: true
    AvailabilityZone:
      Fn::Select:
      - '1'
      - Fn::GetAZs:
```

```
          Ref: AWS::Region
    CidrBlock:
      Ref: PublicSubnet02Block
    VpcId:
      Ref: VPC
    Tags:
    - Key: Name
      Value: !Sub "${AWS::StackName}-PublicSubnet02"
    - Key: kubernetes.io/role/elb
      Value: 1

PrivateSubnet01:
  Type: AWS::EC2::Subnet
  Metadata:
    Comment: Subnet 03
  Properties:
    AvailabilityZone:
      Fn::Select:
      - '0'
      - Fn::GetAZs:
          Ref: AWS::Region
    CidrBlock:
      Ref: PrivateSubnet01Block
    VpcId:
      Ref: VPC
    Tags:
    - Key: Name
      Value: !Sub "${AWS::StackName}-PrivateSubnet01"
    - Key: kubernetes.io/role/internal-elb
      Value: 1
```

```
PrivateSubnet02:
  Type: AWS::EC2::Subnet
  Metadata:
    Comment: Private Subnet 02
  Properties:
    AvailabilityZone:
      Fn::Select:
      - '1'
      - Fn::GetAZs:
          Ref: AWS::Region
    CidrBlock:
      Ref: PrivateSubnet02Block
    VpcId:
      Ref: VPC
    Tags:
    - Key: Name
      Value: !Sub "${AWS::StackName}-PrivateSubnet02"
    - Key: kubernetes.io/role/internal-elb
      Value: 1

PublicSubnet01RouteTableAssociation:
  Type: AWS::EC2::SubnetRouteTableAssociation
```

```
  Properties:
    SubnetId: !Ref PublicSubnet01
    RouteTableId: !Ref PublicRouteTable

PublicSubnet02RouteTableAssociation:
  Type: AWS::EC2::SubnetRouteTableAssociation
  Properties:
    SubnetId: !Ref PublicSubnet02
    RouteTableId: !Ref PublicRouteTable

PrivateSubnet01RouteTableAssociation:
  Type: AWS::EC2::SubnetRouteTableAssociation
  Properties:
    SubnetId: !Ref PrivateSubnet01
    RouteTableId: !Ref PrivateRouteTable01

PrivateSubnet02RouteTableAssociation:
  Type: AWS::EC2::SubnetRouteTableAssociation
  Properties:
    SubnetId: !Ref PrivateSubnet02
    RouteTableId: !Ref PrivateRouteTable02

ControlPlaneSecurityGroup:
  Type: AWS::EC2::SecurityGroup
  Properties:
    GroupDescription: Cluster communication with worker nodes
    VpcId: !Ref VPC
```

```
Outputs:

  SubnetIds:
    Description: Subnets IDs in the VPC
    Value: !Join [ ",", [ !Ref PublicSubnet01, !Ref PublicSubnet02, !Ref
PrivateSubnet01, !Ref PrivateSubnet02 ] ]

  SecurityGroups: eith
    Description: Security group for the cluster control plane communication with
worker nodes
    Value: !Join [ ",", [ !Ref ControlPlaneSecurityGroup ] ]

  VpcId:
    Description: The VPC Id
    Value: !Ref VPC
```

The previous template creates the VPC with two public and two private subnets. One public and one private subnets are deployed to the same Availability Zone. The second public and private subnets are deployed to a second Availability Zone in the same Region. We have chosen this solution because it is that recommended by AWS for critical environment.

After that the VPC is available, it's possible to proceed with the EKS Control plane creation that can be done directly from the AWS console at https://console.aws.amazon.com/eks/home#/clusters through the input of the following parameters:

- Name: INFINITECH-BP
- Kubernetes version: 1.17
- Public and private: true
- role: eksClusterRole(create with in the previous step)
- Encryption: false(Activate if is necessary)
- Subnet: all
- Security groupsInfo: ControlPlane
- Access point Public: True

When the EKS Control plane is available (the creation takes about 15 minutes), it is possible to create the worker nodes.

In order to add the nodes, it is mandatory to create a specific IAM role named WorkerRole with the following permissions:

- AmazonEKSWorkerNodePolicy.
- AmazonEKS_CNI_Policy.
- AmazonEC2ContainerRegistryReadOnly.

Moreover, to facilitate the role management we will add the tag value INFINITECHBPNode.

At this time, it is possible to create from the same AWS console the worker nodes by clicking on the Compute tab and by entering these values:

- Name: T3xlarge_16GB4-4vCPU
- Node IAM role name: WorkerRole
- AMI type: Amazon Linux 2 (AL2_x86_64)
- Instance Type: t3a.xlarge(4vCPU/16GB)
- Disk size: 50 GB
- Minimum size: 2
- Maximum size: 3
- Desired size: 2
- Subnets: private-sub01; private-sub02
- Allow remote access to nodes: true
- Allow remote access from: all

At this stage, the Kubernetes cluster is ready to use.

6.1.3 Namespace, Network and Quote Policies

As previously described, we have decided to implement the INFINITECH Sandbox concept leveraging the Kubernetes namespace feature. Therefore, the first step is to create a dedicated Namespace for each of the INFINITECH target pilots that can be leveraged by each of the owning partners to develop and test their own pilot applications.

To create the namespace we have defined a Kubernetes YAML template like the following:

```
kind: Namespace
apiVersion: v1
metadata:
  name: Pilot1
  labels:
    name: Pilot1
```

Some of these tools (like Jenkins and Gitlab) need to be exposed on the Internet for convenient access during the development and testing phases. All the endpoints will be HTTPS based, with free certificates generated by Let's Encrypt [19], so we do not need to set up a Certification Authority or buy certificates from commercial CAs.

However, we do need public DNS entries on the project domain, mapped to a public IP that exposes our services outside of the Kubernetes cluster.

In order to get the public IP, we will use the AWS ELB (Elastic Load Balancer). This is provisioned automatically when creating a Kubernetes Service with type

LoadBalancer. In our case, this happens while configuring the NGINX Ingress Controller. The NGINX Ingress Controller that will be deployed will have a dedicated namespace named ingress-nginx.

To implement the network policies to isolate each namespace from each other, and eventually also to manage the connection between different PODs within the same namespace (i.e. to guarantee the security requirements), we will implement the Cilium as CNI (Container Network Interface) on EKS.

In order to do that, we will remove the AWS CNI and install the Cilium following these steps:

```
kubectl -n kube-system delete daemonset aws-node
helm repo add cilium https://helm.cilium.io/

helm install cilium cilium/cilium --version 1.7.9 \
  --namespace kube-system \
  --set global.eni=true \
  --set global.egressMasqueradeInterfaces=eth0 \
  --set global.tunnel=disabled \
  --set global.nodeinit.enabled=true
```

After the previous step, it is possible to isolate the namespace implement the following rule (the rules is written in the YAML format):

```
apiVersion: "cilium.io/v2"
kind: CiliumNetworkPolicy
metadata:
  name: "isolate-pilot1"
  namespace: pilot1
spec:
  endpointSelector:
    matchLabels:
      {}
  ingress:
  - fromEndpoints:
    - matchLabels:
        {}
```

Moreover, we will apply the resource quote on memory and CPU that each namespace can consume with this YAML template:

```
apiVersion: v1
kind: ResourceQuota
metadata:
  name: mem-cpu-namespace
spec:
```

```
hard:
  requests.cpu: "800m"
  requests.memory: 2Gi
  limits.cpu: "1"
  limits.memory: 3Gi
```

After the execution of the previous steps, the Deployment view has been completed, as shown in Figure 6.2.

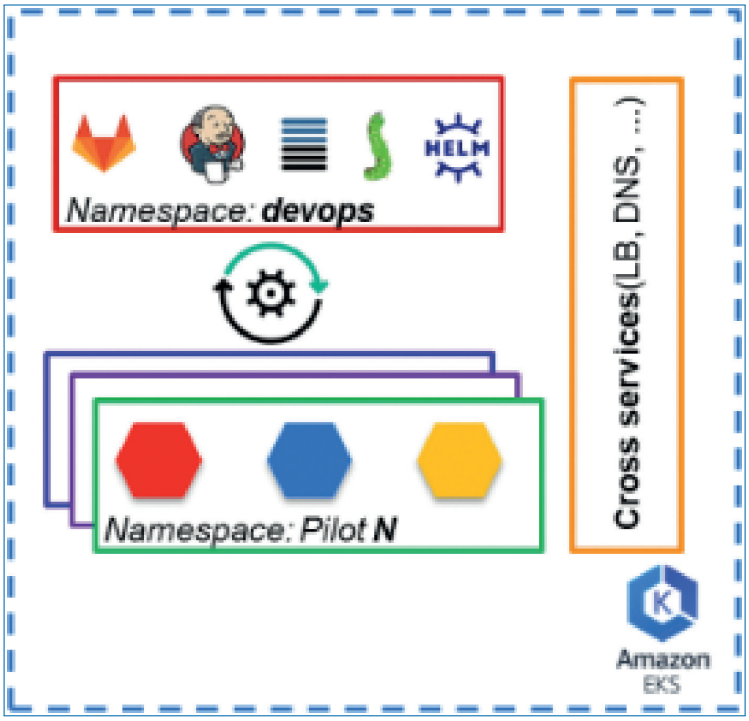


Figure 6.2. INFINITECH blueprint reference testbed.

6.1.4 How to Recreate the Blueprint Testbed for a Specific INFINITECH Pilot

One of the most powerful capabilities and features that are enabled by the technological choices we have made and that are described in the previous sections, is that the blueprint reference testbed that we have created on AWS can be recreated from scratch, with respect to the Deployment view perspective, by each of the partners for their own pilots in two possible ways (see Figure 6.3):

I. On the same AWS cloud provider (a concrete realization is described earlier).
II. In a bare metal environment, leveraging their on-premise private data centre infrastructure or the shared NOVA's Data Centre.

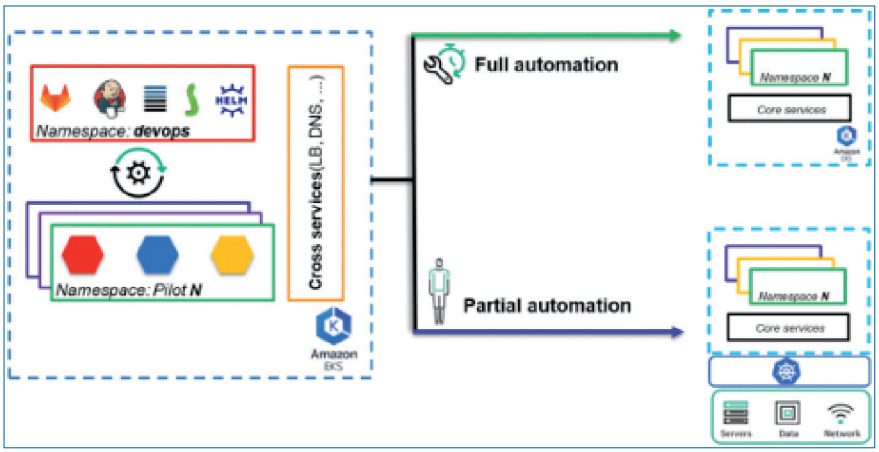


Figure 6.3. Blueprint environment recreation ways.

The two ways are a little bit different each other, because:

I. In the first case, the recreation of the cluster can be done in fully automated way using the eksctl tool provided by AWS.
II. In the second case. it is possible to recreate the cluster in a partially automated way, because in this case as mandatory prerequisite it is necessary to manually create the entire infrastructure that will host the Kubernetes cluster, and only afterwards it is possible to create the cluster using automated tools like Kubespray provided by Kubernetes itself.

Nevertheless, one of the major objectives of having provided a blueprint reference testbed is exactly the powerful concept explained above: a potential easy and straightforward replication of it for all the INFINITECH target pilots' environments.

6.2 Blueprint Environment for Pilot 5b: Bank of Cyprus

As mentioned earlier, as a concrete INFINITECH blueprint environment associated to one of the official INFINITECH pilots, the consortium has selected (at the time of writing) the WP7 Pilot 5b: Business Financial Management (BFM) tools delivering a Smart Business Advise, owned by the partner Bank of Cyprus (BOC), and in particular its initial and preliminary Proof of Concept (PoC) implementation.

This section describes the preliminary concrete realization of such PoC and also setup the basis for the future applications of the concepts described in the previous section for all the INFINITECH target pilots' environments.

6.2.1 Pilot Objectives

Most of today's Financial Management tools for Small Medium Enterprises (SMEs) are geared towards analyzing only past transactions, making such tools inadequate in today's world. Today, SMEs and their customers alike, demand just-in-time processing, transparency and personalized services to assist SME owners not only in understanding better their SME business/financial health but also to be able to decide on the next best action to take.

Thus, Pilot 5b aims to assist SME clients of Bank of Cyprus in managing their financial health in the areas of cash flow management, continuous spending/cost analysis, budgeting, revenue review and VAT provisioning, all by providing a set of AI-powered Business Financial Management tools and harnessing available data to generate personalized business insights and recommendations. Machine learning algorithms, predictive analytics and AI-based interfaces will be utilized to develop a kind of smart virtual advisor with the aim to minimize SME business analysis effort, to focus on growth opportunities and to optimize cash flows performance.

6.2.2 Pilot Workflow

The pilot workflow can be analyzed starting from the datasets that it has to manage. Some of the available datasets require real time data collection, while in others historical data collection is sufficient to provide actionable business insights. In detail, transaction and account data related to the respective SME will be drawn from BOC's repository by a real time/historical data collector as well as transaction and account data from Open Banking (PSD2), as well as BOC's customer data, will utilize a historical data collector. In addition, an external data collector will also be used in order to integrate other related Open Banking/macroeconomic data. The SMEs data source (e.g. ERP/Accounting system) utilization remains optional as consent is required for the collection and processing of such data and its cloud availability being required. Accordingly, it is possible summarize the involved data sources involved in the following list:

- Transaction Data from Open Banking (PSD2).
- Transaction Data from SMEs (optional).
- Other Data (Market).
- Other Data from SMEs (optional).
- Accounts Data from BOC.
- Accounts Data from Open Banking.
- Customer Data from BOC.
- Direct Input from SMEs.

In the target pilot PoC, all data except external macroeconomic data will be pseudonymized (by tokenization) before being uploaded to the IRA.

The cloud Data Repository (within IRA) will then store all collected data (along with the generated insights), past SME financial actions (to measure at what degree the SME actions reflect the recommended insights), as well as minimum user input that is required. A continuous data streaming will connect the Data Repository with the various deployed BFM tools (machine learning algorithms), which would allow the retraining of the respective AI models and the generation of useful insights and recommended actions. A reverse data pseudonymization will then be applied before the processed data move to the bank middleware component that contains composite APIs and produces push notifications, all of which will be offered to the SMEs via Android, iOS and web applications. Upon SME user login, the IRA is also accessed, insights/recommendations picked up from the cloud data repository and provided to the SME user.

The pilot's workflow is depicted in the below Figure 6.4.

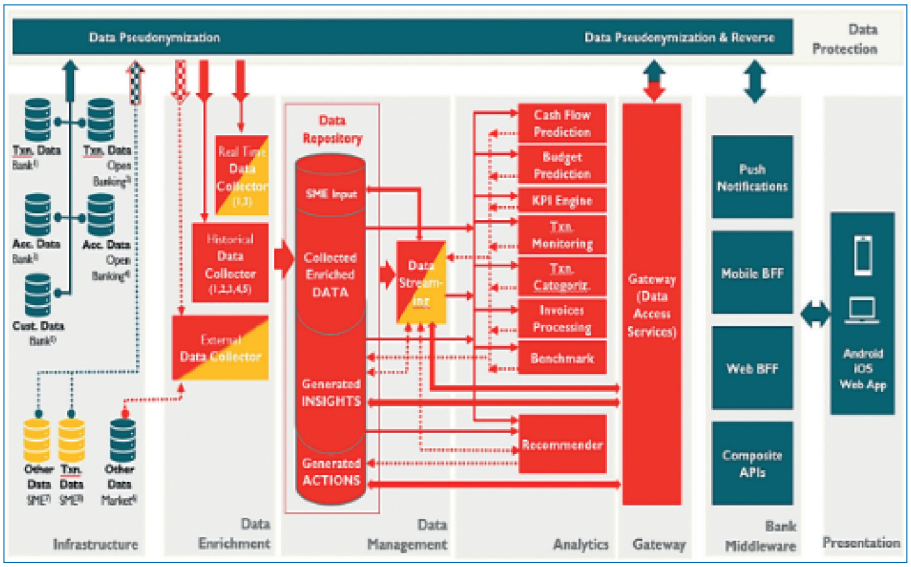


Figure 6.4. Pilot 5B workflow.

6.2.3 Blueprint Reference Testbed Implementation

As stated before, the blueprint reference testbed (Blueprint_v1) refers to the INFINITECH BOC pilot and its initial and preliminary PoC version (Prototype_v1) aims to accommodate the first version of the pilot. Accordingly, the PoC will not use all the available components of the INFINITECH platform, but just a selected subset.

Group	Prototype_v1	Blueprint_v1
On Premise Data source Layer and Data Management	(to be decided)	(to be decided)
Analytics Layer	Cash Flow Prediction ,	Cash Flow Prediction ,
	Transaction Categorization	Transaction Categorization
Data Models and Semantics	Preprocessing	Preprocessing
Data Security and Privacy	-	Data Anonymization
Interfaces	Proxy , API	Proxy , API

Figure 6.5. Components of prototype_v1 and Blueprint_v1 versions.

In general, the blueprint testbed basic assumptions are:

- The pilot does not utilize Blockchain technology, so all the Blockchain-oriented components will not be integrated.
- Data migrated to the blueprint testbed is owned by BOC and will be pseudonymised (using the tokenization technique) before entering the INFINITECH ecosystem. There is no need for an INFINITECH anonymizer for the Prototype_v1 PoC.
- The first PoC version of the pilot (Prototype_v1) may not fully exploit the first version of the pilot Blueprint (Blueprint_v1). There might be additional INFINITECH components (especially regarding the Data Management layer) that remain to be decided.

Prototype_v1 and Blueprint_v1 versions will include the components listed in Figure 6.5.

The next version of the PoC prototype (Prototype_v2) will include stream processing, possibly anonymization and some complementary components of the Analytics group.

The ML/DL models of the Analytics Components (Cash Flow Prediction, Transaction Categorization) will be trained offline and then afterwards loaded to the blueprint analytics component. The models' training process (including the validation and evaluation processes) is depicted in Figure 6.6.

Given the aforementioned points, a REST API will be developed on top, which will deliver real-time information when invoked. This data will be forwarded to the specific analytics component (Cash Flow Prediction, Transaction Categorization) via a REST API/TCP connection or other message broker.

In order to run the pilot inside the blueprint environment a dedicated sandbox named pilot5b has been created. The components deployed in such sandbox are: Analytics, Pre-processing and Anonymizer.

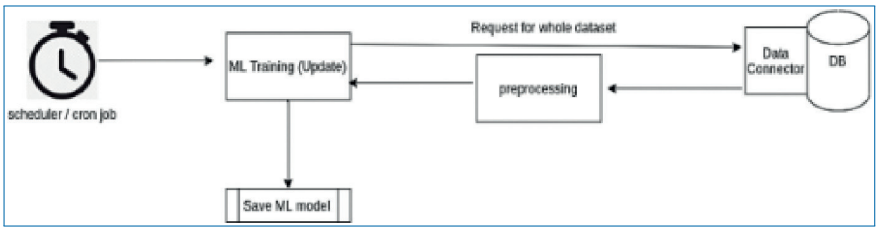


Figure 6.6. Models training process.

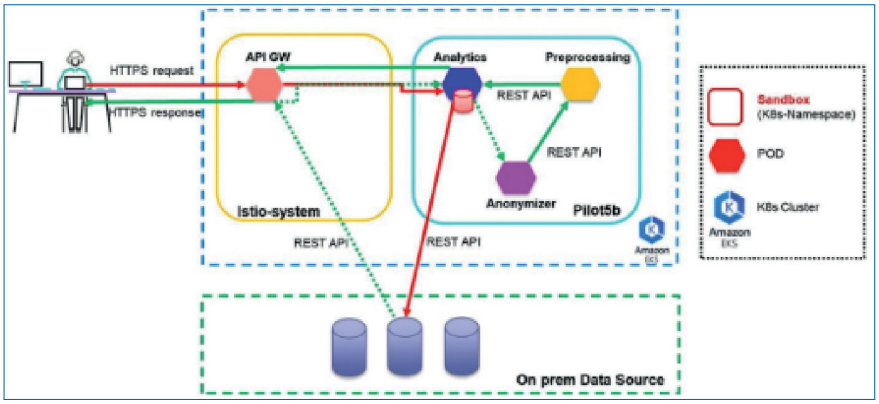


Figure 6.7. Pilot 5b blueprint reference architecture.

The client connections from the external world towards such a sandbox and all sandboxes deployed in the blueprint are managed through an API Gateway (GW), which is a key component of the IRA Interface layer. The API GW in the blueprint environment is based on the Istio [20] software and is deployed in a dedicated sandbox named istio-system.

The process workflow implies that when a user sends a (HTTPS) request it is forwarded by the API GW to the Analytics POD, which will then request the appropriate historical data, stored on premise, from the "Data Management" layer via a data connection/REST API or other message broker. The retrieved data will be first anonymized and then pre-processed based on the approach used for the training procedure. The data will then be forwarded back to the Analytics component in order to be injected in the models to infer the outcome of each model. Finally, the results will be either returned to the user or saved to the internal Data Base, available within the Analytics POD, according to the scope of the process. The proposed blueprint architecture is illustrated in Figure 6.7.

This page intentionally left blank

DOI: 10.1561/9781638282334.ch7

Chapter 7

Lessons Learnt from INFINITECH Pilots

7.1 Lessons Learnt [INFINITECH-Whitepaper-Lessons Learnt]

The context for this section on Lessons Learnt within INFINITECH is the innovations that we have explored in the project (e.g., in architectures, processes, business models), and how they promote the global creation and sharing of knowledge (influenced by prior innovations such as GitHub and GitLab, open science and open Networks of Excellence). Similar work is increasingly commonly used in agile pilots by governments, large corporations and trans-national researchers to create value for society by combining cross-disciplinary expertise as follows:

(a) Insights into ways to create some form of capacity to look forward (e.g., Horizon Scanning), to anticipate and model potentially important changes).
(b) Lessons learned regarding how (and how not to) create some form of analysis of joint experience, the better to look back and to look around.

In Horizon Europe, the combination of (a) and (b) support the spread of expertise in 'Impact' (including disseminating know-what and know-how) and in transferable Pathways (e.g., see cross-cluster, cross-category and cross-program, as in Pathways to Impact.

In INFINITECH, this focus on generalizable and multi-context competence is planned to lead to (c) and (d):

(c) a collection of 'Lessons Learnt' by members of an ecosystem, (e.g., analogous to Open Science, or as in adopting Open Standards for FinTech), leading to:
(d) shareable and generalisable competences in learning to learn, e.g., learning how to refine collections of 'Lessons Learnt'; and how to transfer, generalise, update or correct know-how, to boost community knowledge.

7.2 Cluster/Category 1: Smart, Reliable and Accurate Risk and Scoring Assessment

7.2.1 Pilot #2 – Real-time Risk Assessment in Investment Banking

- **(Near) Real-time risk assessment:** one of the lessons learnt from each pilot (and planned to be quantified in closing months of the project, in terms of the added values of the tool developed within each pilot) is that Pilot #2's findings enable stakeholders'-users, such as traders and risk managers, a near real time risk assessment with several additional futures, including sentiment analysis. The required time to obtain risk estimations is in accordance with the target KPIs. Notably, the VaR/ES of a portfolio consisting of 8 instruments can be obtained in less than 1 second enabling fast what-if analysis. Furthermore, the DeepVaR algorithm – developed in the context of the project – turns out to be a reliable alternative to classical VaR approaches delivering accurate VaR estimations even in periods of high volatility in financial markets. Such enhancements in performance have been validated through the deployment of the P2 in the dedicated testbed and its back-testing in a large amount of historical data.
- **Sentiment analysis:** P2 provides sentiment analysis in financial news using transfer learning based on the FinBERT pre-trained model. The user has access to the original text as well as its sentiment. Thus, they can assess the output of this AI algorithm. The utilisation of transfer learning also facilitates environmental concerns regarding the energy consumption required for the training of AI-based models since the used finbert model is already trained and validated and is used only for inference.
- **Towards a Higher Level of Innovation:** Since many financial institutions still utilise end-of-day (EoD) data rather than intraday data for risk assessment, P2 developed an innovative solution that can be used with real-time data.

- **Highly Interoperable/Scalable:** P2 can be easily incorporated into a variety of data sources and/or systems with little effort. Furthermore, the entire system is very scalable since i) the risk measurements are first calculated in a univariate and parallel manner for each input time series, and ii) its development is built on Kubernetes, which allows distributed computations with resource auto-scaling.
- **Easy-to-understand:** P2 is easy to use and provides a simple visualisation. This signifies that the user does not need any specific expertise. Based on what-if analysis, the tool delivers more information. It can be utilised to get real-time pre-trade risk analysis and can assess them on an individual basis as well as in the context of the entire portfolio.P2 includes the ability to compare the risk assessments with several VaR models (e.g., variance-covariance methods, historical VaR, etc).
- **Stakeholders' Feedback:** Some comments from the stakeholders highlighted the need for more technical details considering ES to be more important than VaR. In this context, P2 assesses the risk of the given portfolios not only in terms of VaR but also in ES terms. It is noted that the novel DeepVaR model provided by the pilot can be used to estimate both risk metrics.

7.2.2 Pilot #15 – Open Inter-banking Pilot

- **Semantic specialisation:** General purpose semantic engines are not effective in understanding banking related concepts. Pilot 15 tries to specialise a state-of-the-art model such as BERT on the banking semantic domain.
- **Efficiency of the training process:** Minimise the training effort as key to ensure a fast development for innovative AI-based solutions. Pilot 15 worked on a weakly-supervised method optimised against legacy semantic resources.
- **Collaborative approach:** Open innovation paradigm is a core value to face the challenges of a complex and fast-growing scenario. Pilot 15 enhanced collaboration within competitors through a shared governance.
- **Research driven approach:** Working in an experimental and competitive environment may help in enhancing flexibility and promote a research driven mindset. Pilot 15 enforced the process of continuous research and evaluation of experimental artefacts.

7.3 Cluster/Category 2: Personalized Retail and Investment Banking Services

INFINITECH Cluster#2 includes four (4) Pilots, whose purpose is to offer custom-made financial services for both retail and investment banking. These are

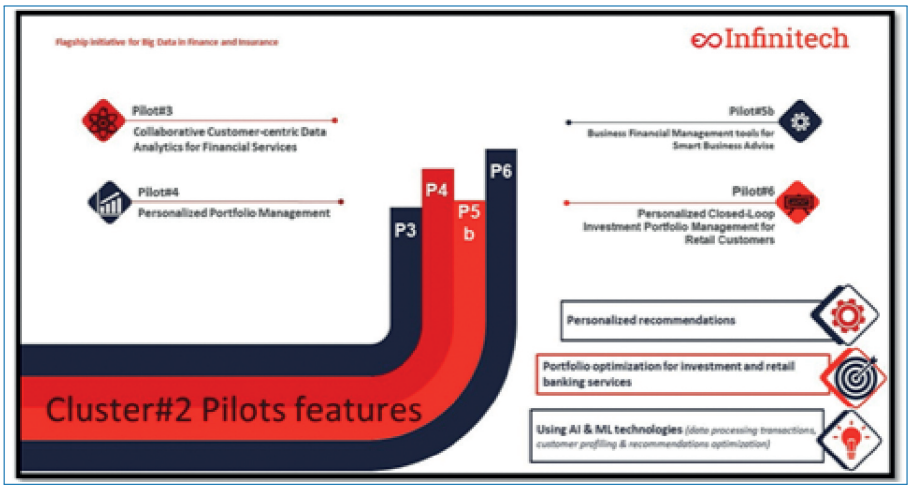


Figure 7.1. Cluster #2 pilots overview.

several personalized amenities built on customer-centric analytics and personalized digital assistants, having as an ultimate objective to deliver improved and/or new technology and business services for banks, SMEs, financial institutions, and governmental agencies, aiming to increase customer experience through personalized proposals and more effective servicing. Despite the fact that they address different financial services, they use mainly AI and ML technologies, such as data processing transactions, customer profiling and recommendations optimization. Cluster #2 Pilots' similarities are driven both by their organizations and customers' needs for personalized offering services as shown in the Figure 7.1.

7.3.1 Pilot #3 – Collaborative Customer-centric Data Analytics for Financial Services

Pilot #3 aims at leveraging the process of verification, named Know Your Customer (KYC), within regulated banking and financial activities, by which secure and trustworthy data sharing and analysis of financial information (customer, account, and transaction data) should be provided. It particularly focuses on identifying unlawful activity, such as money laundering and bank fraud, in relation to financial activities concerning human trafficking, in the context of a global initiative named "Stop the TraffiK" and by the use of Traffik Analysus Hub and Traffik Analysis Hub Ontology (TAHO) and INFINITECH Technologies.

- P#3 addresses KYC issues for banks by using AI to search for Human Trafficking data typologies within the bank's own customer data. KYC process will become more effective if data typologies that give results are shared between

banks. This throws up data privacy issues. EU law on Data Privacy currently can be interpreted as giving criminals data privacy with certain derogations and "let outs" for public interest. Whether a bank's KYC needs can be said to be in the public interest is unclear. On one hand using a negative KYC result to undermine the economics of human trafficking is certainly in the public interest, on the other hand a bank is not a prime engine of law enforcement and as such may not be able to make a clear judgement that their customer is carrying on criminal activities and so is not entitled to data privacy. For banks to be able to share the red flag typologies that have produced positive KYC results, data privacy policy will have to take the P#3 product into account. One member of P#3 is BPFI who can influence Irish and EU DP policy as it applies to banks in Ireland.

- The use of data typologies as the means to sharing across multiple banks allows the red-flag indicators to be shared without actual customer data being exchanged. This allows the participating banks to remain compliant with their GDPR obligations. It should also be noted that the banks will use the outputs from the P#3 service to help identify areas of concern to which the bank fraud teams can apply their focus. There will always be a human involved in the final decision to act on the red-flags. The P#3 service will not be used to automatically decline any customer application or activity.
- The P#3 develoiped the data typologies identification as a tool to make it more more powerful with more participants using it. P#3 is actively disseminating the tool via webinars with the Joint Intelligence Group (JIG) and the International-JIG in April. P#3 also disseminates the use of the tool via a monthly analysts' call of all users of the Traffik Analysis Hub (TAH) (up to 65 organisations). BoI is a known leading contributor of Human Trafficking alerts to law enforcement and has had positive media exposure for doing so. BoI expects the P#3 tool to enhance its already industry-leading stance on human trafficking. These initiatives are expected to add significantly to the number of P#3 users when it is finally launched on the market. P#3 have now engaged a second large bank in the Irish ecosystem to test the service. Between these two banks we can reach over 75% of Ireland's banked population.
- Because of Data privacy concerns regarding data sharing, the P#3 tool will be used within the bank's own data set rather than bank data being shared to the hub where the P#3 tool resides. Even so it is envisaged that P#3 users should share typological patterns that are produced within their datasets – this will be an ongoing tension between the banks sharing patterns and their regulatory obligations.

- Sharing data typologies rather than data will also allow participating banks to edit and evolve existing data typologies. If in using a data typology, one of the banks finds a way to better refine the combination of flags, that bank can then resubmit the enhanced data typology back into the common library. This will allow other banks to more easily ingest the enhancement rather than have to build entire new rules.

7.3.2 Pilot #4 – Personalized Portfolio Management ("Why Private Banking cannot be for Everyone?")

Pilot #4 developed a Portfolio Construction and Optimization algorithm (Prive Optimizer or "AIGO"), as well as improved and expanded its capabilities as an artificial intelligence engine to support better and personalised investment propositions for retail clients. This will enable the approach to make "Private Banking like services available for everyone". By developing a customer-demand and preferences driven service, Pilot #4 provides a leading technology that offers a fully scalable digitised and personalised advisory and wealth management journey for financial institutions (e.g., banks, EAMs, investment firms, insurance companies, and brokerage firms) and market participants.

- Professional wealth management: High human interaction and significant management costs. Pilot 4 illustrates that this process can be digitized to a large extent and that this approach can be extensively modified to individual product variations and investors' needs.
- Technology vendors and financial institutions tend to use terminology that is not familiar to the masses. Pilot 4 tries to use illustrations at various points of user journey to graphically support decisions for a non-financially educated investor.
- Challenge: standardized data, long and correct data histories for each asset and to find the right risk & return balance. A fairly equal distribution of the underlying investment universe is thus pursued. This will make it possible to cover a large range of potential and individual investment preferences.
- Change of Investors' Requirements and Targets: Megatrends, economic and political effects may require regional reassessments or rebalancing of investments. This rebalancing, a repetitive task of checking whether a used portfolio needs to be readjusted, is a permanent effort, which might be fully digitized.
- Transparency as a Trust Building Element: We show volatility data, positive portfolio effects in mixing higher and lower risk investments in the given portfolio, as well as provide portfolio comparison options etc. We thus try to maximize the possibilities for the investor to analyse and understand individualized proposals.

7.3.3 Pilot #5b – Business Financial Management (BFM) Tools Delivering Smart Business Advice

Pilot#5b revolves around SMEs and their needs for business financial management tools. SMEs today are reconsidering their banking relationship and as a result leaving their primary banking provider vulnerable to the increasing competition. Establishing a versatile platform that can unlock new services and can expose & seamlessly integrate evolving technologies represents a key success factor. It offers a Business Financial Management solution (BFM) on top of their core banking activities, aiming at adding real value to the bank's SME customers, as well as attracting new ones through this new offering.

- Data tokenization proved to be a valuable anonymization methodology which reduced compliance/regulatory challenges, in particular relating to GDPR items. It however increased the data preprocessing needs and caused a delay to the introduction of near real-time data streaming functionalities related to transactional data.
- Outsourcing Banking data to cloud providers can turn out to be a quite lengthy and complex undertaking, with no specific guidelines except from a EBA document (EBA Recommendations on outsourcing to cloud service providers, 2017) in place. Having a structured approach and a detailed Testbed Development guideline document at hand is of crucial importance to safeguard the successful cloud implementation.
- Conducting and processing the results of 1:1 SME workshops proved to be a cumbersome and time intensive task, resulting in contradicting preferences and propositions among SMEs, mostly due to the different nature of their core business. Grouping SME per industry could materialize into constructive feedback.
- To attract SME customers, the BFM platform services added value should be clearly communicated and showcased with examples on how the various offerings can support the business owner with everyday financial cash flow tasks so that he can focus on his/her core business.

7.3.4 Pilot #6 – Personalized Closed-Loop Investment Portfolio Management for Retail Customers

Pilot#6 developed a Personalized Investment Recommendations Platform for Retain Customers, leveraging the customers' risk profile and other features. The platform will be available to NBG financial advisors, who not only examine each customer's transactional activity, but also take into account similarities and patterns

among customers, targeting at increasing customer satisfaction, maintaining and reinforcing the customer relationship.

- Extended Clientele: Pilot 6 develops algorithms that aim in Customer Profiling and categorization according to their intention to invest. The above are based not only in questionnaire input but also in transactional activity. The innovation in that part is that the analysis focuses to all Retail Customers and not only to highly affluent.
- UI for Financial Advisor: Pilot 6 wants to create a user friendly UI, so as to illustrate consolidated all the necessary information to Financial Advisors.
- User Friendly UI: Pilot 6, through the UI, wants to provide a "full" image of any personalized proposition to customers.
- Holistic View: A) Enriching the Recommendation Engine and the personalized proposition to the customer, using on top, information from social networks which offers an extra ranking to the proposals.
- Holistic View: Propositions are provided at a significantly lower level compared to other recommendation engines.

7.4 Cluster/Category 3: Predictive Financial Crime and Fraud Detection

INFINITECH Cluster#3 includes five pilot systems that involve Predictive Financial Crime and Fraud Detection. The related pilots intend to provide advanced financial products and services for banks, supervisory authorities, financial institutions and governmental agencies, aiming to prevent and protect against financial crimes and fraudulent activities. The pilots are built on top of a mix of advanced technologies, based on AI and ML, as well as Big Data and Blockchain.

7.4.1 Pilot #7 – Avoiding Financial Crime

Pilot #7 will provide a module for calculating fraud prevention and detection models that help the banks to enhance their current cybersecurity policies and controls to avoid financial crime. By means of unsupervised machine learning and complex modelling, and supported by advanced computational power (near-quantum) technologies, the solution will provide near-real time, operational risk level considering the end-customer's normal behavior, which could greatly improve detection of financial frauds and reduce losses to the banks and society at large.

- AI Assessment:
  - Limited Risk: No assessment of creditworthiness or credit scoring in this system
  - CXB Analysts shall be notified that they interact with an AI system
- Data requirements:
  - Data quality and size of synthetized data needs to be revised regularly
  - Data Governance shall be taken seriously in advance
- AI Model:
  - XGBoost is a favourable AI framework compared to scikit.learn.
  - Integration in the INFINITECH ML/DL library to be discussed
- Deployment:
  - Constraints (data, company policies, etc.) require thorough planning
  - Design needs some scrutinizing and adaptation for onsite requirements at CXB
  - Synthetized data shall be required due to GDPR issues.

7.4.2 Pilot #8 – Platform for Anti Money Laundering Supervision (PAMLS)

The pilot will develop a platform named PAMLS (Platform for AML Supervision) with several tools to enhance and improve the risk-based supervision. Screening tool will process and analyse data from different sources and with methods based on artificial intelligence (AI) and machine learning (ML) techniques, try to recognise unusual patterns and relationships among data, that could indicate typologies and risks of money laundering/terrorist financing (ML/FT) at the level of individual financial institutions (FI). Detected patterns will feed Risk assessment tool, which will assess the FI risk from ML/FT perspective, and this will enable supervisory authority (in case of Pilot #8 BOS) to focus its resources on more high-risk FIs. PAMLS will also have additional functionalities, a Distribution channel that will enable the secure data exchange, and a Search engine: allowing supervisors to look for a specific transaction or a sample of transactions.

- Pilot development takes place in close cooperation between the data provider and the end user (BOS) and technical partner (JSI). To ensure an **agile approach**, regular meetings and workshops with different stakeholders (project team members, supervisors, IT, legal and compliance, technical partner, other experts) were organized. The aforementioned approach has proven to be effective in detecting and addressing identified issues and challenges.

- One of the main challenges was **data quality**. Several iterations were required to prepare data of required quality that were further enriched and pseudo-anonymised. During the data preparation a list of data preparation rules was created.
- Due to **complex legal and compliance requirements**, especially in regard to data protection and ethics, experts from Legal and Compliance departments are included in the project team, to ensure that the development of the pilot is carried out in accordance with legal and ethical requirements.

7.4.3 Pilot #9 – Analyzing Blockchain Transaction Graphs for Fraudulent Activities

Blockchain crypto currencies and tokenized assets that are obtained fraudulently can go through various transfers on the blockchain. Pilot #9 aims to detect such fraudulent activities on massive blockchain transaction graphs. Since blockchain data is constantly accumulating and will be growing at increasing rates. in the future, a parallel scalable transaction graph analysis system is being developed that runs on HPC clusters and that can process the growing transaction graph without encountering performance bottlenecks.

Specifically, Pilot #9 will provide the following product and services:

(i) An open web-based service that operates on massive Ethereum and Bitcoin public blockchain data and reports fraudulent crypto-currency and token transaction activity tracing that is accessible by both common end-users as well as by bigger financial institutions. Free basic service to common end-users as well as paid or agreement based additional customized services will be provided to larger organizations and agencies.

(ii) Token transaction analysis services on the permissioned Hyperledger Fabric which is currently not offered by other companies.

- There are well established competitors: But we are not aware of existing approaches that pay special focus on HPC and scalability. For sustainability of the system, scalability is important which we have learnt to achieve by developing our software using parallel programming.
- Technological development issues: New blockchain transaction throughputs are increasing to thousands of transactions per second which means billions of transactions need to be handled. Our system is ready for such workloads.
- Stakeholders' engagement and stakeholders' feedback: User interface is being improved. National CBDC efforts also need scalable transaction tracing technology.

- Regulatory compliance: Crypto asset regulations are still evolving. Our system can help organizations to meet the new guidelines on crypto assets published by FATF.
- AI assessment: Minimal Risk: We also use other graph algorithms, and all results are checked by human personnel.

7.4.4 Pilot #10 – Real-time Cybersecurity Analytics on Financial Transactions' BigData

The purpose of this service is to enable security-related anomalies to be identified while they are occurring, if possible, by proactively monitoring and taking timely action on such potential security threats. The software component will be able to monitor in real time the financial transactions of a domestic and mobile banking system and will use machine learning models, alongside and in combination with traditional high-efficiency analysis techniques, applied on high-volume real data flows.

Thus, the pilot will move from the current post-event detection approaches to a new real-time approach that will be based on Big Data Analytics (BDA) technologies. For this pilot's reference scenario, the business service to be delivered is reliant on advances in precise and fine-grain financial fraud analysis and detection.

Such a business service will allow to meet two goals:

- The early detection of new and subtle types of frauds. Since fraudsters keep innovating novel ways to scam people and online systems, it becomes crucial to apply AI/ML methods to detect outliers in large transactional datasets and be robust to changing patterns.
- The reduction of the number of false positives which are usually analysed to understand if they are real fraud attempts or not. To this aim, it is very important to be able to train, validate and test ML models to make the most accurate ones operational.
    - Technological capabilities. The system proposed is mainly designed and developed on top of a Data Science and Machine Learning advanced frameworks and open-source technologies for design, deployment, execution, and monitoring of big data analytics workflows.
    - Full potential achievement. The system proposed is based on ML model training by (small) synthetic data. This point was considered a weakness: actual degree of innovation can be assessed only with real data processing. Pilot#10 has been validated on big data got by an AI-based Synthetic Transactions Generator.
    - Scalability and Integration. The proposed solution is built referencing to a cloud based architecture able to scale computing and storage resources

thanks to a workflow orchestration engine that leverages the capabilities of Kubernetes for cloud resource management.

- Trustworthiness and compliance. The system proposed is designed by taking into account the banking data sovereignty. Such a prototype will be released to be integrated on-premise so to keep all the existing internal procedures and avoid the external data processing.

7.4.5 Pilot #16 – Data Analytics Platform to Detect Payments Anomalies Linked to Money Laundering Events

INFINITECH Pilot #16 "Data Analytics Platform to detect payments anomalies linked to money laundering events" aims to build a data analytics platform to help Nexi AML team to discover, monitor and analyze suspicious scenarios related to money laundering through digital card payments, using machine learning and advanced analytics methods. To build the data platform, the partners are collecting historical data from clients in an anonymized format. The anonymization phase is performed before source data is uploaded into the Data Lake, where the processing steps are performed. This allows to ensure the data confidentiality and privacy of customers. Since all data is anonymized, we can use it directly to perform data processing steps and the training of machine learning algorithms.

During the pilot development, we encountered different types of challenges and experiences. They gave us the opportunity to work on novel approaches, data, applications, errors, and improvement opportunities. We summarized them in terms of:

- Type: project level involved. E.g.: technical, organizational, etc.
- Change/Performance Consequence: task/issue requiring actions.
- Cause: what led to the change/consequence.
- Action: what we did to achieve the goal.

Type	Change or Performance Consequence	Cause	Action
Technical Back-end	Initial difficulties in using Neo4J	Neo4J is a novel graph database approach, especially applied to AML data	Frequent interaction to share preliminary testing results and increase collective knowledge
Technical Back-end	Data QA	Assessment performed on edges and nodes rather than tabular data	Training analysis on known use-cases to establish benchmarks for further progressive evaluations

Type	Change or Performance Consequence	Cause	Action
Communication	Slight mismatches on mock-up characteristics	Insufficient communication on progression of features characteristics	Adoption of Agile methodology approach and sprint reviews
Stakeholders involvement	Suitability of internal AML stakeholders involvement with daily working tasks	AML stakeholders need to get used to a novel approach while performing daily activities	Development of uses-cases on AML subjects and/or merchants which actually are under investigation
Technical Infrastructure	Deployment of pilot components on INFINITECH infrastructure	Integration of Nexi infrastructure and deliverables with INFINITECH testbed	Development of a preliminary flowchart to organize and prioritize the development and deployment of each component

7.5 Cluster/Category 4: Personalized Usage-Based Insurance (UBI) Pilots

The concept of Usage Based Insurance (UBI) is taken to the next level within this cluster by adding AI technologies to the real world data collected from the users' environment. Here, different IoT infrastructures are exploited as novel data sources and combined with different ML/DL technologies to define, develop and train specific AI models. These AI powered models will assist users' classification and risk detection processes, and, in turn, support customised services offered to insured parties and insurance companies. Specifically, this cluster develops a pilot focused on connected cars and motor insurance and another pilot related to activity trackers and e-health (See Figure 7.2).

7.5.1 Pilot #11 – Personalised Insurance Products Based on IoT Connected Vehicles

This pilot is oriented to the car insurance business, relying on rising connected car infrastructures and considering each connected vehicle as an IoT entity. Big Data, HPC and AI technologies will be applied here, together with the new Vehicle to Infrastructure (V2I) paradigm, to identify and define diverse driving profiles and so, classify real drivers according to their behaviour. This new business innovation revisits the way insurance premiums are calculated and so, supports a new set of insurance products adapted to the car insured clients (Figure 7.3).

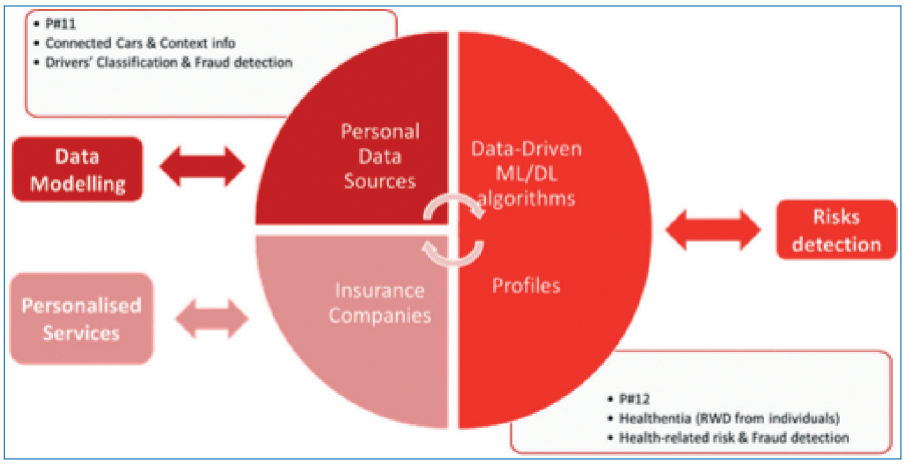


Figure 7.2. Personalised insurance pilots' common approach.

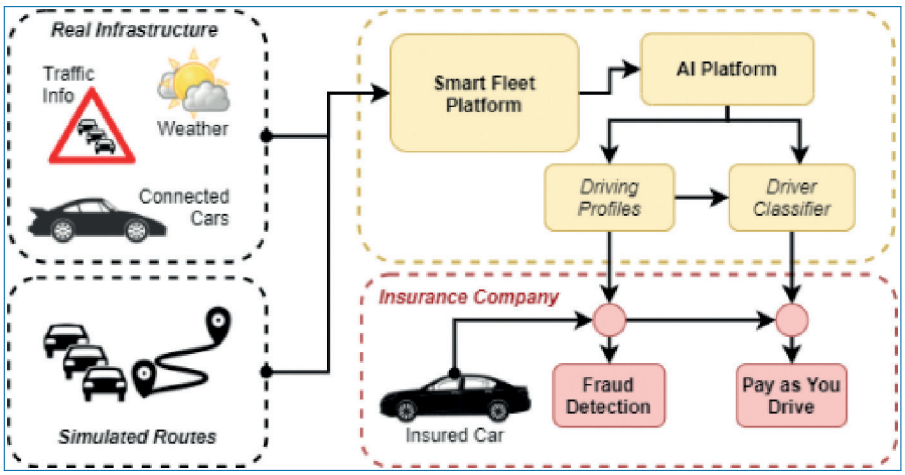


Figure 7.3. Personalised motor insurance pilot architecture. UBI paradigm implementation.

During the deployment of this motor insurance pilot we've collected a set of lessons learnt, based on the experience of our internal insurance company, our data providers and the stakeholders' meetings we organised. From different points of view, this feedback will help on the future expansion of the proposed business model.

From (Motor) Drivers' perspective, engagement of drivers to provide their technical vehicle data on a voluntary basis is fundamental:

- from volunteers'/drivers' perspective, to train and test the AI models. Without any initiative and benefits derived from this data gathering process, it is

likely to have dropouts or not use their vehicles on a regular basis. Within the pilot, our vehicles' provider engaged real drivers.

- from insureds'/drivers' perspective, to evaluate their driving behaviour. In this real scenario, it would be necessary to find ways (even if there is premiums' increase) to ensure this participation. Finding different ways to capture technical vehicle datasets is another challenge.

From Insurance companies' perspective, these are carefully studying the different possibilities that the AI insights are providing. They're waiting to see the final services working to apply the aforementioned insights. We need to remark the relevance of the AI applied here and to show some previous results. This must be done to evolve the final business model.

From data gathering process and homogenisation point of view, technical vehicle datasets, even based on OBD standards, depend a lot on the manufacturer and may differ, in terms of volume and measurement units, from one deployment to another. Context information (weather, traffic, etc.) is also quite dependent on the scenario. Not all scenarios can provide the same set of context data. Moreover, context information need to be relevant to the areas of driving.

From ML engineers' perspective, data used for the AI system is biassed by the set of real drivers involved (in terms of profiles and area of driving) so the AI prototypes would be really relevant for these set of users.

7.5.2 Pilot #12 – Real World Data for Novel Health-Insurance Products

The pilot focuses on health insurance and analyses the impact of continuously monitored Real-World Data (RWD), captured from users' smart devices (bracelets, smartphones, etc.) or personalized risk assessment. The continuous personalized risk assessment is offered via analytics presented in personal and cohort-based dashboards, as well as on predictions of developed AI models. Once assessed, the risk facilitates customization of health insurance products.

- Insured users' perspective
  - End users' questionnaire findings
    - They are not really willing to share information with their insurance company, regardless of possible financial benefits. They are more willing to share lifestyle data, but by no means any medical data.
    - Personal health risk analysis is wanted as part of their benefit in the program: We are employing a virtual coach to give them feedback & advice.

- Analysis of pilot study feedback
  - The mobile application is really important and constitutes the main reason for pilot study dropouts.
  - They want less text (legal text is needed though, in accordance to GDPR and Medical Device Regulation)
  - They feel the app is too demanding on their phones (less processing power requirements)
  - They wish for more feedback (see point on risk analysis reaching the users from the workshop above)
  - They ask for freedom on reports (to be able to report what they want, not what we are interested in). It is tricky though to utilise such unstructured reports in model learning.
  - The app was considered as easy to use by the end users.
  - The added value of the app was low or not clear to the end users. They want to have an app that differentiates from the standard health app on their mobile phones.
  - They want to have more physical activities included within the app, for example: biking and swimming.
  - They feel that adding health data manually takes too much time, for example in the case of the number of glasses of water they drank during that day. They do not remember how much they drank when they open the app, and they do not think about it just after drinking a glass.

- Overall take-away messages:
  - Life insurance is a program that runs for life! Gamification features are needed to keep the attention of users, ensuring compliance. Especially so since the users are not under strict doctor's orders.
  - The functionalities of the Healthentia app, focus mainly on measuring health data. An explanation for the high number of drop-outs we experienced could be that for mHealth users, it is not about collecting health data. There needs to be an added value to be willing to use the Healthentia app for a long-term period. So, even though potential clients can use the overview of their health data to monitor their own health as a first step in living a healthy life, this is not enough to keep engaging them in the long-term.

- Insurance companies' perspective (feedback from stakeholders' workshops)
  - They acknowledge the need for premium adaptation
  - They perceive the importance of monitoring dashboards higher than model recommendations

- Regulators' perspective (feedback from regulators' workshop)
  - They see regulatory implications in both data collection and the application of the data for continuous risk assessment. As far as the data collection is concerned, applications that do so are commonplace. As to the health insurance use case, there are now the first such programs offered to the public, rewarding "good behaviour" with discounts.
  - It is the need to comply with regulation that clashes with some users' requests for less and simpler text in the agreements.
- ML engineers' perspective
  - Data quality (adherence, diversity or volume) is VERY important to obtain functional models.
  - Different devices provide different quality of data (accuracy, timmings, etc.) and also the correct usage of these devices by the insured user is something that must be taken into consideration.
  - Anonymization is mandatory regulation-wise. Unfortunately the assessment of its effect on model quality was not possible in the study carried out. The data quality was low to begin with, leading to models without acceptable performance. This seems to be equally true for anonymised and non-anonymised data alike.
  - Model bias is a huge problem. The population the model is to be applied to is everyone, not a small segment. Learning such a model would require an unmanageable number of study participants giving information for too long a time. The only alternative is to start such programs without the AI part, and build the necessary models per population segment over the years.

7.6 Cluster/Category 5: Configurable and Personalized Insurance Products for SMEs and Agro-Insurance

Cluster 5 is composed of two pilots that base their analysis on big data from different sources, both open and from satellite imagery to gather real-world and real-time data. The objective is to develop AI powered services to enhance risk profiling. These pilots will develop their own architecture by combining INFINITECH and pilot's specific technologies, configure their corresponding sandboxes and run their testbeds, all within the INFINITECH framework.

Cluster 5 pilots are intended to provide configurable and personalised insurance products based on alternative data sources and big data, including various personalised services based on customer centric analytics and personalised digital assistants (Figure 7.4).

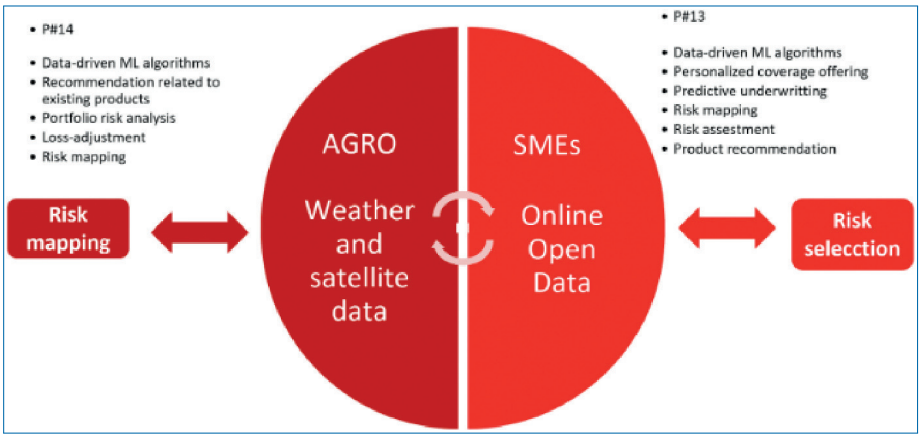


Figure 7.4. Cluster 5 layout.

7.6.1 PIlot #13 – Alternative/Automated Insurance Risk Selection – Product Recommendation for SME

Focuses by obtaining the data in open sources and the application of machine learning, the pilot will be able to monitor the changes in the risks, so we will be able to radically improve the risk management that companies face in the development of their daily activity.

The activities carried out based on the collection of comments and opinions by Pilot #13 regarding stakeholders have been done through a workshop and the completion of online questionnaires. The online questionnaires had to be carried out as the organized workshop was not sufficiently attended and we considered the reinforcement of feedback through the questionnaires.

Therefore, one of the lessons learned is that mass workshops involving different Pilots, even if they are insurance Pilots, are not effective. This is because, given the stage of maturity of the project, we believe that it is more effective for each Pilot to seek personalized feedback from the stakeholders who can contribute most to the project. Stakeholders, for example from an insurance company, are very different whether we are talking about business, agriculture, auto or health insurance.

That said, the feedback we got was very interesting and enriching as it proved that Pilot #13 is considered a high priority by the stakeholders, immediately applicable and highly innovative both technologically and business-wise.

In addition to the above with respect to the stakeholders, we complement the lessons learned in the following areas:

- Data – The density of the raw information obtained from open sources has been much more complex than anticipated but two measures have been taken

during implementation which have proved effective. The incorporation of other sources to allow for overlap and complementarity of data.

- Technology – It is not viable, when using the data, to depend only on obtaining it from sources via API, so the technology of Wenalyze's concurrent autonomous micro robots has been adapted to this reality in order to, as mentioned in the previous point, improve the density and quality of the data.
- Market and insurers – It is undeniable that a data-driven industry such as insurance needs the solution developed in pilot #13, however the go to market to interact with insurers has taken longer than expected. They are not certain of the need for such data, so the best way we have found to demonstrate the need for the solution is to run pilots and demos with their data to compare the different qualities of the data.

7.6.2 Pilot #14 – Big Data and IoT for the Agricultural Insurance Industry

Provide Insurance companies with a robust and cost-effective toolbox of functions and services – allowing them to alleviate the effect of weather uncertainty when estimating risk of AgI products, reduce the number of on-site visits for claim verification, reduce operational and administrative costs for monitoring of insured indexes and contract handling, and design more accurate and personalized insurance/coverage products.

Pilot actors – AgI companies and stakeholders involved in the Insuretech workshop series, had significantly contributed with their comments and suggestions to what resulted to be the INFINITECH AgI toolbox MVP.

Having engaged a large number of actors coming from different regions, providing different coverages/AgI insurance products, showcasing different levels of IT infrastructure, we witnessed a large diversification in the needs priority but at the same time a large homogenization of expectations. This was more or less what we expected since we started the process of involving AgI actors coming from different enterprises (maturity of markets and sectors, different sizes, varied market outreach, significantly different portfolio of products etc).

Weather information (past climatic and now-casting) seems to be a priority for the majority of the AgI actors compared to the use of satellite data; AgI actors readiness to utilise upscaled weather information whether in operations management or in risking assessment and premium underwriting, is higher. On the other hand although the use of satellite data for event and contract related information (affected area, severity of impact, prioritisation of in-field visits etc) was also medium to high, though AgI actors readiness to utilise was hindered by current regulation

restrictions (majority of Agriculture Insurance systems require the in-field assessment of damage).

Despite these variations and respective legislation "obstacles", the INFINITECH-AgI toolbox is being praised for its functionalities and easiness of use, its cost-efficiency and availability of services, enabling a holistic approach to AgI operations.

DOI: 10.1561/9781638282334.ch8

Chapter 8

Conclusions

8.1 Conclusions [D8.4]

We begin the book by providing a background on the FinTech. The FinTech signifies an on-going technological revolution in financial sector. For example, means of payment is now shifted towards payment by contact-less mechanism or cards, replacing old means of payment such as cheques. According to Juniper FinTech is: *The use of technology to underpin the delivery of financial services*. FinTech has two main pillars, namely technology and financial system. FinTech disrupts the financial services industry by enhancing customer experience, increasing the speed of service, and reducing operating cost through digitization. Despite resistance to changing environment, financial and insurance services sectors are now embracing this disruption. The waves of digitization, Financial Technology (FinTech) and Insurance Technology (InsuranceTech) are rapidly transforming the financial and insurance services industry. Data is epi·centre of these changes. The vast majority of digital transformation applications for the finance and insurance sectors are data intensive. Available bigdata sources fules new more automated, personalized, and accurate services.

This book provide a summary of INFINITECH Reference Architecture (RA) as the core element of INFINITECH Way Foundation. Reference Architecture (RA) of the INFINITECH project aimed to develop smart, autonomous and

personalized services in the European fnance and insurance services ecosystem. The INFINITECH-RA will specify a set of building blocks that will support advanced BigData, AI and IoT applications. These building blocks will support scalable, unified, and interoperable data collection from different sources and databases. We provide a detail description of RA in BIGDATA/IOT/AI in finance and insurance sectors. INFINITECH-RA will specify the structuring principles that will drive the integration of these building blocks in real-life solutions.

The book explains innovative technologies for financial sectors applied in INFINITECH project. The INFINITECH project rely in the capabilities of its partner members to produce value that is exploitable beyond the lab or Proof of Concept of ideas. The expertise from academia and research converges with industrial products to provide solutions to existing problems in financial sector. The existing capabilities of data management in the industry is not sufficient. INFINITECH will provide solutions for integrated data management over the wide range of databases and data sources used by BigData, IoT and AI applications in finance/insurance.

The marketplace is considered the main driver for assets utilization among members of the INFINITECH ecosystem, as well as the main artefact for exploitation and sustainability of the project. As the specifications and detailed architecture of the INFINITECH marketplace have been provided in past deliverables, and in specific D8.1 and D8.2, this document described and analyzed, in greater detail compared to the previous version of the current deliverable (i.e. D8.3), the final version of the developed REST API endpoints of the INFINITECH Marketplace's back-end.

The marketplace offers ready-to-use solutions covering a wide range of modern business and technical needs, focusing, as most pilots of INFINITECH project, on the Finance and Insurance sectors. Of course, most assets are related to Big Data and AI techniques and algorithms, with a variety of datasets, experimentation results and models being already available at the marketplace. Other than ready-to-use algorithms, frameworks and combined solutions, some assets offer state-of-the-art IoT and Blockchain solutions.

Additionally, the INFINITECH Marketplace aspires to become a prominent multi-sided market platform and Virtualized Digital Innovation Hub (VDIH) and continue adding value to practitioners, organizations and communities of the Finance and Insurance sector long after the INFINITECH Project has been completed.

Finally, this book and the progress presented within, is well aligned with the project's goals and overall strategy, awaiting additional content to be populated from the various events and third-party activities (e.g., Hackathons, Webinars, etc.), that

will be the content of the last WP8 deliverables (related to third parties), until the end of the project.

The project's validated solutions at the technological or business level are made available at the project multi-sided market platform (marketplace) and/or a VDIH for wider use and commercial exploitation. Stakeholders of the digital finance/insurance and FinTech/InsuranceTech ecosystem interact with INFINITECH through market platform and the VDIH. It is a public web-based environment with various APIs, able to store several types of assets that may derive/result from the separate procedures and mechanisms that are either implemented in the scope of the project or not. We illustrate the component of marketplace such as back-end, frond-end, and VDIH. The back-end is the main component of the marketplace. It consists of three different layers and implements the main functionalities for the assets management. The front-end is the fourth layer of the market platform. It is a web-based server that presents the offered assets to the users, with a friendly UI. The front-end converts all interfaces of the back-end (REST API) into user friendly interfaces and provides automated forms and processes that make it easier for users to interact with the back-end and benefit from its stored assets. The VDIH organized into Training Activities and Innovation Services, while the Training Activities include courses, workshops and webinars, the Innovation Services includes acceleration programs. The VDIH pages provide all the resources available on the platform related to that type of content.

To create community around the INFINITECH Marketplace and enrich it with new information, the new functionalities such as social login and add new information were added to the marketplace. Social login fulfil one of the objectives of the INFINITECH wich is to create a digital finance ecosystem of innovation, with IoT, Blockchain, BigData and AI solutions and services. To enrich the digital finance ecosystem of innovation, it was created forms to give the users the opportunity to share their solutions and services on the INFINITECH Marketplace, because it is important that information continues to evolve and increase. Moreover, we describe marketplace usage scenarios. The two scenarios include upload a VDIH and consult a VDIH. We complete the section by providing an overview of the baseline technologies, interfaces, API related to assets, users and descriptions, search functionalities, validation, upload description, upload and retrieve assets, and authentication scenarios. A detailed list of assets uploaded to marketplace is also provided.

We describe the tools and techniques that will be leveraged to implement the testbeds and sandboxes concepts within the INFINITECH project, considering that the INFINITECH-RA is designed leveraging a paradigm based on a microservices architecture implementation, with services interacting among them through REST APIs. Containers benefits, microservices approach, Kubernetes

containers orchestration, Kubernetes architecture, INFINITECH testbeds, and INFINITECH sandboxes are detailed and explained in a concise manner.

We follow on to delineate tools and techniques for management of dataset. This chapter describes the tools and techniques that will be leveraged for the management of datasets within the INFINITECH project and how they are linked and mapped with the concepts and techniques related to testbeds and sandboxes and with respect to the blueprint reference testbed environment. We illustrate data sources and data access and Datasets management from the blueprint reference environment perspective.

Moreover, we provide an overview of the INFINITECH blueprint reference testbeds. This chapter describes the initial design and implementation of the INFINITECH blueprint reference testbed, through the actual realization (with a full compliance) of the INFINITECH-RA development and deployment views, in terms of the concrete specification and realization of the fundamental and target INFINITECH concepts of testbeds, sandboxes and datasets management, and related tools and techniques for their effective setup and deployment in the INFINITECH pilots and validation scenarios. We discuss development view, creation of the EKS INFINITECH cluster, namespace, network and quote policies, and H=how to recreate the blueprint testbed for a specific INFINITECH Pilot.

Last, we examine lesson learnt within each pilot. The context for this section within INFINITECH is the innovations that we have explored in the project (e.g., in architectures, processes, business models), and how they promote the global creation and sharing of knowledge (influenced by prior innovations such as GitHub and GitLab, open science and open Networks of Excellence). Similar work is increasingly commonly used in agile pilots by governments, large corporations, and trans-national researchers to create value for society by combining cross-disciplinary expertise.

DOI: 10.1561/9781638282334.ch9

Chapter 9

References

9.1 Infinitech Public Deliverables

D1.3 Data Management Plan
D1.5 Data Management Plan – II

D2.1 User Stories and Stakeholders' Requirements – I
D2.2 User Stories and Stakeholders' Requirements – II
D2.3 Reference Scenarios and Use Cases – I
D2.4 Reference Scenarios and Use Cases – II
D2.5 Specifications of INFINITECH Technologies – I
D2.6 Specifications of INFINITECH Technologies – II
D2.7 Security and Regulatory Compliance Specifications – I
D2.8 Security and Regulatory Compliance Specifications – II
D2.9 Initial Specification of Testbeds, Data Assets and APIs – I
D2.10 Initial Specification of Testbeds, Data Assets and APIs – II
D2.11 Data Model Specifications – I
D2.12 Data Model Specifications – II
D2.13 INFINITECH Reference Architecture – I
D2.14 INFINITECH Reference Architecture – II
D2.15 INFINITECH Reference Architecture – III

D3.1 Hybrid Transactional/Analytics Processing for Finance and Insurance Applications – I
D3.2 Hybrid Transactional/Analytics Processing for Finance and Insurance Applications – II
D3.3 Hybrid Transactional/Analytics Processing for Finance and Insurance Applications – III
D3.4 Integrated (Polyglot) Persistence – I
D3.5 Integrated (Polyglot) Persistence – II
D3.6 Data Streaming and Data at Rest Queries Integration – I
D3.7 Data Streaming and Data at Rest Queries Integration – II
D3.8 Data Streaming and Data at Rest Queries Integration – III
D3.9 Automatic Parallelization of Data Streams and Intelligent Pipelining – I
D3.10 Automatic Parallelization of Data Streams and Intelligent Pipelining – II
D3.11 Automatic Parallelization of Data Streams and Intelligent Pipelining – III
D3.12 Data Governance Framework and Tools – I
D3.13 Data Governance Framework and Tools – II
D3.14 Data Governance Framework and Tools – III
D3.15 Regulatory Compliance Tools – I
D3.16 Regulatory Compliance Tools – II
D3.17 Regulatory Compliance Tools – III
D3.18 Integrated (Polyglot) Persistence – III

D4.1 Semantic Models and Ontologies – I
D4.2 Semantic Models and Ontologies – II
D4.3 Semantic Models and Ontologies – III
D4.4 Semantic Streams Analytics Engine – I
D4.5 Semantic Streams Analytics Engine – II
D4.6 Semantic Streams Analytics Engine – III
D4.7 Permissioned Blockchain for Finance and Insurance – I
D4.8 Permissioned Blockchain for Finance and Insurance – II
D4.9 Permissioned Blockchain for Finance and Insurance – III
D4.10 Blockchain Tokenization and Smart Contracts – I
D4.11 Blockchain Tokenization and Smart Contracts – II
D4.12 Blockchain Tokenization and Smart Contracts – III
D4.13 Encrypted Data Querying and Personal Data Market – I
D4.14 Encrypted Data Querying and Personal Data Market – II
D4.15 Encrypted Data Querying and Personal Data Market – III

D4.16 Visualization Frond-End for Aggregated Information – I
D4.17 Visualization Frond-End for Aggregated Information – II

D5.1 Library of Parallelized Incremental Analytics – I
D5.2 Library of Parallelized Incremental Analytics – II
D5.3 Library of Parallelized Incremental Analytics – III
D5.4 Framework for Declarative and Configurable Analytics – I
D5.5 Framework for Declarative and Configurable Analytics – II
D5.6 Framework for Declarative and Configurable Analytics – III
D5.7 Library of ML/DL Algorithms – I
D5.8 Library of ML/DL Algorithms – II
D5.10 Data Management Workbench and Open APIs – I
D5.11 Data Management Workbench and Open APIs – II
D5.12 Data Management Workbench and Open APIs – III
D5.13 Datasets for Algorithms Training & Evaluation – I
D5.14 Datasets for Algorithms Training & Evaluation – II

D6.1 Testbeds Status and Upgrades – I
D6.2 Testbeds Status and Upgrades – II
D6.3 Testbeds Status and Upgrades – III
D6.4 Tools and Techniques for Tailored Sandboxes and Management of Datasets – I
D6.5 Tools and Techniques for Tailored Sandboxes and Management of Datasets – II
D6.6 Tools and Techniques for Tailored Sandboxes and Management of Datasets – III
D6.7 Sandboxes in Incumbent Testbeds – I
D6.8 Sandboxes in Incumbent Testbeds – II
D6.9 Sandboxes in Incumbent Testbeds – III
D6.10 Sandboxes for FinTech/InsuranceTech Innovators – I
D6.11 Sandboxes for FinTech/InsuranceTech Innovators – II
D6.12 Sandboxes for FinTech/InsuranceTech Innovators – III
D6.13 Testbeds Support and Certification Services – I
D6.14 Testbeds Support and Certification Services – II
D6.15 Testbeds Support and Certification Services – III

D7.1 Report on Pilot Sites Preparation – I
D7.2 Report on Pilot Sites Preparation – II
D7.18 Pilot Sites Synergies and Collaboration – I
D7.19 Pilot Sites Synergies and Collaboration – II

D7.20 Pilots' Evaluation and Stakeholders' Feedback – I
D7.21 Pilots' Evaluation and Stakeholders' Feedback – II
D7.22 Pilots' Evaluation and Stakeholders' Feedback – III

D8.1 Market Platform and VDIH Specifications – I
D8.2 Market Platform and VDIH Specifications – II
D8.3 BigData and AI Solutions Marketplace – I
D8.4 BigData and AI Solutions Marketplace – II
D8.5 IoT and Blockchain Solutions Marketplace – I
D8.6 IoT and Blockchain Solutions Marketplace – II
D8.7 Virtualized Digital Innovation Hub – I
D8.8 Virtualized Digital Innovation Hub – II
D8.9 Third-Party FinTech and InsuranceTech Solutions – I
D8.10 Third-Party FinTech and InsuranceTech Solutions – II
D8.11 Ecosystem Support Services and Third Party Solutions – I
D8.12 Ecosystem Support Services and Third Party Solutions – II

D9.1 Detailed Dissemination/Communication Plan
D9.2 Dissemination and Communication Activities – I
D9.3 Dissemination and Communication Activities – II
D9.4 Dissemination and Communication Activities – III
D9.5 Dissemination and Communication Activities – IV
D9.6 Contributions to Standards, Associations and Clusters – I
D9.7 Contributions to Standards, Associations and Clusters – II
D9.8 Contributions to Standards, Associations and Clusters – III
D9.9 Community Building Report – I
D9.10 Community Building Report – II
D9.11 Community Building Report – III
D9.12 Exploitation and Sustainability Plan – I

9.2 References

Accenture – Technology Vision for Insurance 2021 https://www.accenture.com/us-en/insights/insurance/technology-vision-insurance-2021.

A. Narayanan and V. Shmatikov, "De-anonymizing Social Networks," in 2009 30th IEEE Symposium on Security and Privacy, May 2009, pp. 173–187.

A. Machanavajjhala, D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber, "Privacy: Theory meets Practice on the Map," 2008 IEEE 24th International Conference on Data Engineering. 2008, doi: 10.1109/icde.2008.4497436.

A. Crespo, "PowerPoint presentation: eIDAS-Compliant Cross-Border Authentication – Alberto Crespo." 2018.

"ARIES – ReliAble euRopean Identity EcoSystem." [Online]. Available: https://www.aries-project.eu/. [Accessed: 07-Jul-2021].

A-System Requirements. ICAO – Security and facilitation. Retrieved June 24, 2021, from https://www.icao.int/Security/FAL/PKD/BVRT/Pages/System-Requirements.aspx.

"Article 29 Working Party. Opinion 05/2014 on Anonymisation Techniques (2014)." https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf (accessed Sep. 14, 2020).

A. Narayanan and V. Shmatikov, "De-anonymizing Social Networks," in 2009 30th IEEE Symposium on Security and Privacy, May 2009, pp. 173–187.

A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," ACM Trans. Knowl. Discov. Data, vol. 1, no. 1, p. 3–es, 2007.

"Amazon Kinesis.", https://aws.amazon.com/kinesis/.

Acerbi, C., & Tasche, D. (2002). Expected Shortfall: a natural coherent alternative to Value at Risk. Economic notes, 31(2), 379–388. Economic notes.

Allen, D. M. (2012). The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction. Technometrics, 16(1), 125–127. doi/abs/10.1080/00401706.1974.10489157.

Alpaydin, E. (2014). Introduction to Machine Learning (3rd ed.). MIT Press.

Altman, N.S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46(3), 175–185.

A. W. Service, "Amazon Kubernetes Service," [Online]. Available: https://aws.amazon.com/it/quickstart/architecture/amazon-eks/.

A. C. P. C. J. T. E. B. a. R. S. W. Radomski, "ERC-1155 multi token standard," 2015. [Online]. Available: https://eips.ethereum.org/EIPS/eip-1155. [Accessed 20 November 2021].

Alberto Crespo García, Nicolás Notario McDonnell, Carmela Troncoso et al., Privacy- and Security-by-Design Methodology Handbook, December 2015,

https://fr.slideshare.net/richard.claassens/priparemethodologyhandbookfinalfeb242016 and https://ipen.trialog.com/images/ipen/a/a1/PRIPARE_Methodology_Handbook_Final_Feb_24_2016.pdf (accessed in November 2020).

ARIES consortium, "D4.1 ARIES prototype instantiation." 2018, [Online]. Available: https://www.ariesproject.eu/sites/default/files/aries/public/contentfiles/deliverables/D4.1%20ARIES%20prototype%20instantiation.pdf.

A. Polyviou, P. Velanas, and J. Soldatos, "Blockchain Technology: Financial Sector Applications Beyond Cryptocurrencies," Proceedings, vol. 28, no. 1, p. 7, 2019.

A. Camerra, J. Shieh, T. Palpanas, T. Rakthanmanon, and E. J. Keogh. Beyond one billion time series: indexing and mining very large time series collections with iSAX2+. Knowledge and Information Systems (KAIS), 39(1):123–151, 2014.

A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," ACM Trans. Knowl. Discov. Data, vol. 1, no. 1, p. 3–es, 2007.

Alberto Crespo García, Nicolás Notario McDonnell, Carmela Troncoso et al., Privacy- and Security-by-Design Methodology Handbook, December 2015, https://fr.slideshare.net/richard.claassens/priparemethodologyhandbookfinalfeb242016 and https://ipen.trialog.com/images/ipen/a/a1/PRIPARE_Methodology_Handbook_Final_Feb_24_2016.pdf (accessed in November 2020).

AWS, "Amazon EKS," Amazon, 2020. [Online]. Available: https://aws.amazon.com/eks/?nc1=h_ls.

A. Z. R. a. B. S. Dandekar, "Comparative evaluation of synthetic data generation methods," in ACM Conference (Deep Learning Security Workshop), 2017.

A. Palavalli, D. Karri, and S. Pasupuleti, "Semantic internet of things," in 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), 2016, pp. 91–95.

A. Narayanan and V. Shmatikov, "Robust De-anonymization of Large Sparse Datasets," in 2008 IEEE Symposium on Security and Privacy (sp 2008), May 2008, pp. 111–125.

Alexandrov, A., Benidis, K., Bohlke-Schneider, M., Flunkert, V., Gasthaus, J., Januschowski, T., Maddix, D.C., Rangapuram, S., Salinas, D., Schulz, J., Stella, L., Caner Turkmen, A., and Wang, Y, (2019). GluonTS: probabilistic time series models in Python. arXiv preprint arXiv:1906.05264.

"Apache Calcite", https://calcite.apache.org/.

"Article 29 Working Party. Opinion 05/2014 on Anonymisation Techniques (2014)." https://ec.europa.eu/justice/article-29/documentation/opinionrecommendation/files/2014/wp216_en.pdf (accessed Sep. 14, 2020).

AWS, "What is aws," Amazon, 2020. [Online]. Available: https://aws.amazon.com/it/what-is-aws/.

Apache.org, "Maven," [Online]. Available: https://maven.apache.org/.
"Apache Tomcat," [Online]. Available: http://tomcat.apache.org/. [Accessed 15 October 2020].
Abouzeid, K. Badja-Pawlikowski, D. Abadi, A. Silberschatz, A. Rasin, "HadoopDB: an architectural hybrid of MapReduce and DBMS technologies for analytical workloads", PVLDB, vol. 2, pp. 922–933 (2009).
Apache Impala, http://impala.apache.org/.
Apache Drill – Schema-free SQL Query Engine for Hadoop, NoSQL and Cloud Storage, https://drill.apache.org/.
A. De Nicola and M. Missikoff, "A lightweight methodology for rapid ontology engineering," Communications of the ACM, vol. 59, no. 3, pp. 79–86, 2016.
A. Martidis, I. Tomasic, and P. Funk, "Deliverable 5.3: CREATE Interoperability." Aug. 25, 2014.
Adar, E., and Huberman, B. (2001). A market for secrets. First Monday, 6(8).
A. Mueen, Y. Zhu, M. Yeh, K. Kamgar, K. Viswanathan, C. Gupta, and E. Keogh. The fastest similarity search algorithm for time series subsequences under euclidean distance, August 2017. http://www.cs.unm.edu/~mueen/FastestSimilaritySearch.html.
Alexis Hope, Catherine D'Ignazio, Josephine Hoy, Rebecca Michelson, Jennifer Roberts, Kate Krontiris, and Ethan Zuckerman. (2019). Hackathons as Participatory Design: Iterating Feminist Utopias. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, 61.
Arnab Nandi and Meris Mandernach. (2016). Hackathons as an informal learning platform. In Proceedings of the 47th ACM Technical Symposium on Computing Science Education. ACM, 346–351.
Angarita, M. A. M., & Nolte, A. (2020). What do we know about hackathon outcomes and how to support them?–A systematic literature review. In International Conference on Collaboration Technologies and Social Computing (pp. 50–64). Springer, Cham.
Anti-money laundering and countering the financing of terrorism legislative package. (2021). European Commission. https://ec.europa.eu/info/publications/210720-anti-money-laundering-countering-financing-terrorism_en (Accessed: January 2022).
AIOTI |The Alliance for the Internet of Things Innovation. Retrieved January, 2022 from https://aioti.eu/https://aioti.eu/.
https://cordis.europa.eu/project/id/951972.
Apache SparkTM – Unified Analytics Engine for Big Data. (2019). Apache.org. https://spark.apache.org/ (Accessed: January 2022).
Apache Kafka. https://kafka.apache.org/ (Accessed: January 2022).

Apache Flink: Stateful Computations over Data Streams. (2014). Apache.org. https://flink.apache.org/ (Accessed: January 2022).

"Apache Storm.", http://storm.apache.org/.

Alexander Nolte, Ei Pa Pa Pe-Than, Anna Filippova, Christian Bird, and Herbsleb James D. Scallen, Steve. (2018). You Hacked and Now What? – Exploring Outcomes of a Corporate Hackathon. Proceedings of the ACM on Human-Computer Interaction 2, CSCW (2018), 129:1–129:23.

Anne-Laure Mention (2019) The Future of Fintech, Research-Technology Management, 62:4, 59–63, DOI: 10.1080/08956308.2019.1613123.

Arasu A. et al. (2016) STREAM: The Stanford Data Stream Management System. In: Garofalakis M., Gehrke J., Rastogi R. (eds) Data Stream Management. Data-Centric Systems and Applications. Springer, Berlin, Heidelberg.

A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," IEEE Pervasive Comput., vol. 2, no. 1, pp. 46–55, Jan. 2003.

A. Simitsis, K. Wilkinson, M. Castellanos, U. Dayal, "Optimizing analytic data flows for multiple execution engines", in ACM SIGMOD, 2012, pp. 829–840.

A. Tomasic, L. Raschid, P. Valduriez, "Scaling access to heterogeneous data sources with DISCO", IEEE Trans. On Knowledge and Data Engineering, vol. 10, pp. 808–823, 1998.

B. Gedik and L. Liu, "Location Privacy in Mobile Systems: A Personalized Anonymization Model," 25th IEEE International Conference on Distributed Computing Systems (ICDCS'05). doi: 10.1109/icdcs.2005.48.

Baik Hoh and M. Gruteser, "Protecting Location Privacy Through Path Confusion," in First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05), Sep. 2005, pp. 194–205.

Baik Hoh, M. Gruteser, Hui Xiong, and A. Alrabady, "Enhancing Security and Privacy in Traffic-Monitoring Systems," IEEE Pervasive Comput., vol. 5, no. 4, pp. 38–46, Oct. 2006.

"Beating financial crime: Commission overhauls anti-money laundering and countering the financing of terrorism rules. Press Release. Website of the European Union." 2021. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_3690 [Accessed July-2021].

Baptista, A. I., & Silva, A. P. (2018). FINANCIAL TECHNOLOGIES EFFECT ON FINANCIAL SERVICES FROM AN OPEN INNOVATION PERSPECTIVE. Theory and Applications in Game Theory, 179.

B Campbell C Mortimore, "Security Assertion Markup Language (SAML) 2.0 Profile for OAuth 2.0 Client Authentication and Authorization Grants." 2015, [Online]. Available: https://tools.ietf.org/html/rfc7522.

Brown, R. G. (1959). Statistical forecasting for inventory control. McGraw/Hill.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.

Becker and Whisler, 1967 https://econpapers.repec.org/article/ucpjnlbus/v_3a40_3ay_3a1967_3ap_3a462.html.

Busilytics – Digital Transformation https://www.busilytics.com/digital-transformation/.

"Balisage Paper: RESTful Service Description Language (RSDL)," [Online]. Available: http://www.balisage.net/Proceedings/vol10/html/Robie01/BalisageVol10-Robie01.html. [Accessed 24 03 2021].

BOOST4.0, D2.5 – BOOST 4.0 Reference Architecture Specification v1, Available at: https://cordis.europa.eu/project/id/780732/results [Accessed: 15-May-2020].

Big Data Value Association. BVA SRIA—European big data value strategic research and innovation agenda. 2017. http://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf. Accessed 18 Feb 2020.

Boid D, Chang W. NIST Big Data Interoperability Framework: Volume 6, RA Version 2. NIST Big Data Program. 2018. https://bigdatawg.nist.gov/_uploadfiles/NIST.SP.1500-6r1.pdf.

Bons, R.W.H., Alt, R., Lee, H.G. et al. Banking in the Internet and mobile era. Electron Markets 22, 197–202 (2012). https://doi.org/10.1007/s12525-012-0110-6.

B. Kolev, P. Valduriez, C. Bondiombouy, R. Jiménez-Peris, R. Pau, J. Pereira, "CloudMd-sQL: querying heterogeneous cloud data stores with a common language", Distributed and Parallel Databases, vol. 34, pp. 463–503. Springer (2015).

Botts, M., Percivall, G., Reed, C. and Davidson. J. OGC Sensor Web Enablement: Overview and High Level Architecture. Technical report, OGC, December 2007.

Barros, R., & Santos, S. (2018). A large-scale comparison of concept drift detectors. Information Sciences, 451–452, 348–370.

Bhattacharya, A. (2020) Effective Approaches for Time Series Anomaly Detection. towardsdatascience.com.

Bergstra, J., & Bengio, Y. (2012) Random Search for Hyper-Parameter Optimization. Journal of Machine Learning Research, 13(1), 281–305.

Bifet, A., & Gavaldà, R. (2007). Learning from time-changing data with adaptive windowing. In Proceedings of the Seventh SIAM International Conference on Data Mining, SDM '07, SIAM, pp. 443–448.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. Springer Link. doi.org/10.1023/A:1010933404324.

Buchanan, B. G. (2019). Artificial Intelligence in finance. The Alan Turing Institute, 00(00), 3–49. 10.5281/zenodo.2626454.

"Blockchain basics: Hyperledger Fabric," IBM Developer, 2020. [Online]. Available: https://developer.ibm.com/technologies/blockchain/articles/blockchain-basics-hyperledgerfabric/. [Accessed 29 August 2020].

Bishop, C.M. (2006). Pattern recognition. Machine learning 128, no. 9.

Breiman, L. (2001). Random Forests. Machine Learning, 45 (1), pp. 5–32.

Berg, Larry K., et al. "Evaluation of a modified scheme for shallow convection: Implementation of CuP and case studies." Monthly Weather Review 141.1 (2013): 134–147.

Briscoe G, Mulligan C, (2014). "Digital innovation: the hackathon phenomenon" Creative works London/QMUL, London. Available at: http://www.creativeworkslondon.org.uk/wpcontent/uploads/2013/11/Digital-Innovation-TheHackathon-Phenomenon1.pdf.

Bard Rosell, Shiven Kumar, and John Shepherd. (2014). Unleashing innovation through internal hackathons. In Innovations in Technology Conference (InnoTek), 2014 IEEE. IEEE, 1–8.

BDVA Task Force 6: Technical. https://www.bdva.eu/task-force-6 (Accessed: January 2022).

B. Schölkopf, J. C. Platt, J. C. Shawe-Taylor, A. J. Smola, and R. C. Williamson, Estimating the Support of a High-Dimensional Distribution. Neural Comput. 13, 7, 1443–1471, 2001.

BDVA: "BDV SRIA – European Big Data Value Strategic Research and Innovation Agenda", Version 4.0, October 2017.

"CEF Digital Home." https://ec.europa.eu/cefdigital/wiki/cefdigital/wiki/display/CEFDIGITAL/CEF+Digital+Home (accessed Sep. 14, 2020).

C. Gómez, "eID under eIDAS Building trust in a digital society – DG CONNECT-European Commission." [Online]. Available: http://st.fbk.eu/sites/st.fbk.eu/files/20180316_eidas_oauth_security_workshop.pdf.

C. Dwork, A. Roth, and Others, "The algorithmic foundations of differential privacy," Found. Trends Theor. Comput. Sci., vol. 9, no. 3–4, pp. 211–407, 2014.

C. Bettini, X. S. Wang, and S. Jajodia, "Protecting Privacy Against Location-Based Personal Identification," in Secure Data Management, 2005, pp. 185–199.

C. A. Ardagna, M. Cremonini, E. Damiani, S. D. C. di Vimercati, and P. Samarati, "Location Privacy Protection Through Obfuscation-Based Techniques," Data and Applications Security XXI. pp. 47–60, 2007, doi: 10.1007/978-3-540-73538-0_4.

Çetintemel U. et al. (2016) The Aurora and Borealis Stream Processing Engines. In: Garofalakis M., Gehrke J., Rastogi R. (eds) Data Stream Management. Data-Centric Systems and Applications. Springer, Berlin.

Chira Barua et al., "The last pit stop? Time for bold late-cycle moves – McKinsey Global Banking Annual Review", McKinsey & Co, 2019.

Compliance Maturity Level https://www.leancompliance.ca/post/capabilities-maturity-model-for-compliance*. (https://www.leancompliance.ca/post/capabilities-maturity-model-for-compliance).

"Consul," [Online]. Available: https://www.consul.io/. [Accessed 10 October 2020].

C. Bondiombouy, B. Kolev, O. Levchenko, P. Valduriez, "Multistore big data integration with CloudMdsQL", Transactions on Large-Scale Data and Knowledge-Centered Systems (TLDKS), pp. 48–74. Springer (2016).

C. Iso and others, "ISO/IEC 2382-1: 1993 Information Technology-Vocabulary-Part 1: Fundamental terms," 1993.

Compton, M., Barnaghi, P., Bermudez, L., Castro, R. G., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., Huang, V., Janowicz, K. Kelsey, W. D., Phuoc, D. L., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheath, A. and Taylor, K. The SSN Ontology of the Semantic Sensor Networks Incubator Group. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, ISSN 1570-8268, Elsevier, 2012.

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. ICML '06: Proceedings of the 23rd International Conference on Machine learning, 161–168. ACM Digital Library. doi.org/10.1145/1143844.1143865.

Chawla, N. V., Herrera, F., Garcia, S., & Fernandez, A. (2018) SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. Journal of Artificial Intelligence Research, 61, 863–905.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. Springer. doi.org/10.1007/BF00994018.

Cristianini, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press. doi.org/10.1017/CBO9780511801389.

"Consent Receipt Specification – Kantara Initiative," Kantara Initiative, 2020. [Online]. Available: https://kantarainitiative.org/download/7902/. [Accessed 1 September 2020].

Cranor C.D., Johnson T., Spatscheck O. (2016) Stream Processing Techniques for Network Management. In: Garofalakis M., Gehrke J., Rastogi R. (eds)

Data Stream Management. Data-Centric Systems and Applications. Springer, Berlin, Heidelberg C.

C. Dwork, A. Roth, and Others, "The algorithmic foundations of differential privacy," Found. Trends Theor. Comput. Sci., vol. 9, no. 3–4, pp. 211–407, 2014.

C. Faloutsos, M. Ranganathan, and Y. Manolopoulos. Fast subsequence matching in timeseries databases. In Proceedings of the International Conference on Management of Data (SIGMOD), pages 419–429, 1994.

CNIL, Infographic on DPIA methodology, retrieved on Mai 13th 2020 via https://www.cnil.fr/sites/default/files/atoms/files/171019_fiche_risque_en_cmjk.pdf.

Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 591–594. [MIRAB12] Microsoft Industry Reference Architecture for Banking (MIRA-B), Microsoft Corporation Whitepaper, May 2012.

"Credit Card Fraud Detection", 2016. [Online]. Available: https://www.kaggle.com/mlg-ulb/creditcardfraud

«Cloning,» [Online]. Available: https://sdorsett.github.io/post/2018-12-24-using-terraform-toclone-a-virtual-machine-on-vsphere/).

«CNCF,» Linux Foundation, [Online]. Available: https://www.cncf.io/.

Christoersen, P.F.: Evaluating interval forecasts. International economic review pp. 841–862 (1998).

Chesbrough, H. (2019). Open innovation results: Going beyond the hype and getting down to business. Oxford University Press.

Chainalysis Team Indirect Exposure: Why you need to look beyond direct counterparties to understand cryptocurrency address risk, https://blog.chainalysis.com/reports/cryptocurrency-risk-blockchain-analysisindirect-exposure, November 2020.

Chesbrough, H., and Bogers, M. (2014) Explicating open innovation: Clarifying an emerging paradigm for understanding innovation, In H. Chesbrough, W. Vanhaverbeke, & J. West (Eds.), New Frontiers in Open Innovation: 3–28. Oxford: Oxford University Press.

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL Towards better implementation of the EU's anti-money laundering and countering the financing of terrorism frameworkCOM/2019/360 final https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=CELEX%3A52019DC0360 (Accessed: January 2022).

"Circuit Breaker Pattern," 2020. [Online]. Available: https://docs.microsoft.com/enus/azure/architecture/patterns/circuit-breaker. [Accessed 10 October 2020].

"Communication in a microservice architecture," [Online]. Available: https://docs.microsoft.com/enus/dotnet/architecture/microservices/architect-micros

ervice-container-applications/communication-inmicroservice-architecture. [Accessed 24 03 2021].

Contentcal.io what-is-a-good-social-media-engagement (https://www.contentcal.io/blog/what-is-a-good-social-media-engagement-rate/#:~:text=A%20good%20engagement%20rate%20for%20Linkedin%20is%20about%202%25%2C%20however,type%20of%20content%20you%20share.)

Context Information Management (CIM); NGSI-LD API Disclaimer. (n.d.). https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.05.01_60/gs_CIM009v010501p.pdf (Accessed: January 2022).

Cilium. [Online]. Available: https://cilium.io/.

Data Protection, Privacy and Electronic Communications (Amendments etc) (EU Exit) Regulations 2019, https://www.scl.org/news/10384-data-protection-privacy-and-electronic-communicationsamendments-etc-eu-exit-regulations-2019 [accessed March 2022].

D. G. Murray, F. McSherry, R. Isaacs, M. Isard, P. Barham, and M. Abadi, "Naiad: A Timely Dataflow System," in Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles, New York, NY, USA, 2013, pp. 439–455.

D. J. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring High-Level Behavior from Low-Level Sensors," in UbiComp 2003: Ubiquitous Computing, 2003, pp. 73–89.

D. Taibi, V. Lenarduzzi and C. Pahl, "Architectural patterns for microservices: a systematic mapping study," 2018.

Docker, "What container," 2020. [Online]. Available: https://www.docker.com/resources/what-container.

D. G. Neil MacDonald, "12 things to get right for successful devsecops," Gartner, December 2019. [Online]. Available: https://www.gartner.com/en/documents/3978490/12-things-to-get-right-forsuccessful-devsecops.

Devlin, J., Chang, M.-W., Lee K., Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), pp. 4171–4186.

D. J. Abadi et al., "The Design of the Borealis Stream Processing Engine," p. 13.

Dealroom.co – The state of European Insurtech 2021 https://dealroom.co/blog/the-state-of-european-insurtech-2021.

D. Yaga, P. Mell, N. Roby and K. Scarfone, Blockchain technology overview, National Institute of Standards and Technology, 2018.

D. Deuber, B. Magri and S. A. K. Thyagarajan, "Redactable Blockchain in the Permissionless Setting," 2019 IEEE Symposium on Security and Privacy (SP), pp. 124–138, 2019.

De Montjoye, Y.-A., Shmueli, E., Wang, S.S., and Pentland, A. (2014). OpenPDS: Protecting the privacy of metadata through SafeAnswers. PloS One, 9 (7), e98790.

Dubey, A., and Pentland, A. (2020a). Differentially-private federated linear bandits. In Proceedings of the Conference on Neural Information Processing Systems (NeurIPS 2020).

Dubey, A., and Pentland, A. (2020b). Kernel methods for cooperative multi-agent contextual bandits. In Proceedings of the 37th International Conference on Machine Learning (ICML), 119:2740–2750.

Dubey, A., and Pentland, A. (2020c). Private and byzantine-proof cooperative decision-making. In Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2020), 357–365.

D. S. E. E. a. N. S. W. Entriken, "EIP-721: ERC-721 non-fungible token standard," 2018. [Online]. Available: https://eips.ethereum.org/EIPS/eip-721. [Accessed 15 November 2021].

Deloitte, "Open banking – Privacy at the epicentre," Deloitte, 2018.

De Marchi, S., Schaback, R., Wendland, H. (2005). Near-optimal data-independent point locations for radial basis function interpolation, Adv. Comput. Math. 3, 23, 317–330.

D. DeWitt, A. Halverson, R. Nehme, S. Shankar, J. Aguilar-Saborit, A. Avanes, M. Flasza, J. Gramling, "Split query processing in Polybase", in ACM SIGMOD, pp. 1255–1266 (2013).

Dr. Ann Cavoukian, former Information and Privacy Commissioner of Ontario, Canada; for PbD see https://iapp.org/resources/article/privacy-by-design-the-7-foundational-principles/ (accessed in November 2020).

DBeaver, SQL UI client, https://dbeaver.io/.

Dimitrios Salampasis and Anne-Laure Mention (2021) Transformation Dynamics in FinTech: An Open Innovation Ecosystem Outlook. https://doi.org/10.1142/12355.

Dong-Hee Shin, Hongbum Kim, Junseok Hwang, Standardization revisited: A critical literature review on standards and innovation, Computer Standards & Interfaces, Volume 38, 2015, Pages 152–157, ISSN 0920-5489, https://doi.org/10.1016/j.csi.2014.09.002.

Directive (EU) 2015/849 of the European Parliament and of the Council of 20 May 2015 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, amending Regulation (EU) No 648/2012 of the European Parliament and of the Council, and repealing Directive 2005/60/EC of the European Parliament and of the Council and Commission Directive 2006/70/EC (Text with EEA relevance) 4th AML

Directive-4AMLD https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32015L0849 (Accessed: January 2022).

Directive (EU) 2018/843 of the European Parliament and of the Council of 30 May 2018 amending Directive (EU) 2015/849 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, and amending Directives 2009/138/EC and 2013/36/EU (Text with EEA relevance) Directive (EU) 2018/843 (Accessed: January 2022).

Debezium https://debezium.io/ (Accessed: January 2022).

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

"DevTest and DevOps for microservices," Microsoft, 2020. [Online]. Available: https://docs.microsoft.com/en-us/azure/architecture/solution-ideas/articles/dev-test-microservice.

De Filippi, P., and McCarthy, S. (2012). Cloud computing: Centralization and data sovereignty- European Journal of Law and Technology, 3(2), available at SSRN: https://ssrn.com/abstract=2167372.

D. Kyriazis, et al, "BigDataStack: A Holistic Data-Driven Stack for Big Data Applications and Operations", BigData Congress, San Francisco, CA, USA, 2018: 237–241.

Document Object Identifier Accessible here: http://www.doi.org/handbook_2000/DOIHandbook-v4-4.1.pdf.

Exxact Corporation. (2020). PyTorch vs TensorFlow in 2020: What You Should Know About These Frameworks. Exxact. https://blog.exxactcorp.com/pytorch-vs-tensorflow-in-2020-what-you-should-know-about-these-frameworks/.

"EUR-Lex – 32014R0910 – EN – EUR-Lex." https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=uriserv:OJ.L_.2014.257.01.0073.01.ENG (accessed Sep. 14, 2020).

EU General Data Protection Regulation (EU-GDPR), https://www.privacyregulation.eu/en/index.htm [accessed March 2022].

European Commission. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da. https://eur-lex.europa.eu/eli/reg/2016/679/oj.

European Commission, "Technology readiness levels (TRL" [Online]. Available: https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf.

EUR-Lex – 32014R0910 – EN – EUR-Lex." https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=uriserv:OJ.L_.2014.257.01.0073.01.ENG (accessed Sep. 14, 2020).

European Commission (2020) On Artificial Intelligence – A European approach to excellence and trust (White paper), COM(2020) 65 final, available at https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

European Banking Authority (2020), EBA report on Big Data and Advanced Analytics, EBA/REP/2020/01, available at https://eba.europa.eu/sites/default/documents/files/document_library//Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf.

E. Union, "Official Journal of the European Union," 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN. [Accessed 05 November 2021].

E. &. Y. G. Limited., "How banks can balance GDPR and PSD2," 2019. [Online]. Available: https://www.ey.com/en_lu/banking-capital-markets/how-banks-can-balance-gdpr-and-psd2. [Accessed 14 November 2021].

EY Global Wealth Research Report. EYG no. 003263-21Gbl. (2021) "Where will wealth take clients next?" Source: https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/wealth-and-asset-management/ey-2021-global-wealth-research-report.pdf.

EY – Global Consumer Banking Survey 2016 – EY – Global. (2016). http://www.ey.com/gl/en/industries/financial-services/banking---capital-markets/ey-global-consumer-banking-survey-2016.

Ernst & Young. 2017. EY FinTech Adoption Index: Fintech Services Poised for Mainstream Adoption in the US With 1 in 3 Digitally Active Consumers Using Fintech. Press release, June 28. https://www.ey.com/us/en/newsroom/news-releases/news-ey-fintech-adoptionindex.

Emily Porter, Chris Bopp, Elizabeth Gerber, and Amy Voida. (2017). Reappropriating Hackathons: The Production Work of the CHI4Good Day of Service. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, 810–814.

Erman, C., (2017). Financial technologies effect on financial services from an open innovation perspective.

ETSI GS CIM 009 V1.5.1 (2021-11) Context Information Management (CIM); NGSI-LD API https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.05.01_60/gs_CIM009v010501p.pdf.

Ethics guidelines for trustworthy AI. (2019, April 7). European Commission. Retrieved January 2022, from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

European Commission. (2020). White Paper on Artificial Intelligence: a European approach to excellence and trust (White Paper No. COM(2020) 65 final). European Commission. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approachexcellence-and-trust_en.

EBA report identifies key challenges in the roll out of Big Data and Advanced Analytics |European Banking Authority (europa.eu) https://www.eba.europa.eu/eba-report-identifies-key-challenges-roll-out-big-data-and-advanced-analytics (Accessed: January 2022).

Ethereum Magicians forum https://ethereum-magicians.org/ (Accessed: January 2022).

Ethereum Trust Alliance Https://Ethtrust.Org/ (Accessed: January 2022).

Ethereum Improvement Proposals, https://eips.ethereum.org/ (Accessed: January 2022).

European Blockchain Services Infrastructure (EBSI), https://ec.europa.eu/digital-single-market/en/european-blockchain-services-infrastructure (Accessed: January 2022).

EBSI Documentation, Ledger Protocols, https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/EBSI (Accessed: January 2022).

Einpresswire – Fintech market boosted by rising demand for digitization in organizations https://www.einnews.com/pr_news/549826454/fintech-market-boosted-by-rising-demand-for-digitization-in-organizations.

ESCP Business school repository https://www-emerald-com.revproxy.escpeurope.eu/insight/content/doi/10.1108/MF-02-2017-0028/full/pdf?title=electronic-finance-recent-developments.

European Commission (2020) On Artificial Intelligence – A European approach to excellence and trust (White paper), COM(2020) 65 final, available at https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligencefeb2020_en.pdf.

European Banking Authority (2020), EBA report on Big Data and Advanced Analytics, EBA/REP/2020/01, available at https://eba.europa.eu/sites/default/documents/files/document_library//Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf.

ESCP Business school repository https://web-a-ebscohost-com.revproxy.escpeurope.eu/ehost/pdfviewer/pdfviewer?vid=4&sid=a9e463e0-f2fd-42c2-bfad-5444579e0422%40sdc-v-sessmgr03.

Emerald Insight – Detecting money laundering transactions with machine learning https://www.emerald.com/insight/content/doi/10.1108/JMLC-07-2019-0055/full/html.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the econometric society, 987–1007.

"ETSI NGSI-LD specification," [Online]. Available: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.01.01_60/gs_CIM009v010101p.pdf.

EY – Four ways COVID-19 is reshaping consumer banking behavior https://www.ey.com/en_kr/banking-capital-markets/four-ways-covid-19-is-reshaping-consumer-banking-behavior.

Electronic Privacy Information Center, Privacy and Human Rights 2005: An International Survey of Privacy Laws and Developments. 2007.

European Union Agency for Cybersecurity, "Pseudonymisation Data protection GDPR Pseudonymisation techniques and best practices," 2019, doi: 10.2824/247711.

ETSI, "ETSI TS 186 020 V3.1.1 – Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); IMS-based IPTV interoperability test specification," Technical, 2011.

EICTA, "EICTA INTEROPERABILITY WHITE PAPER." 2004, [Online]. Available: http://www.agoria.be/www.wsc/webextra/prg/nwAttach?vWebSessionID=8284&vUserID=999999&appl=Agoriav2&enewsdetid=73029&attach=DescFile10.720108001095774198.pdf.

"EU data protection rules," European Commission, 2020. [Online]. Available: https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules_en. [Accessed 27 August 2020].

EPCglobal: Reader Protocol Standard, Version 1.1, 3 Ratified Standard, 4 June 21, 2006.

EPCglobal: The Application Level Events (ALE) Specification, Version 1.1.1 Part I: Core Specification, EPCglobal Ratified Standard, 13 March 2009.

EPCglobal: The EPCglobal Architecture Framework, EPCglobal Final Version 1.2 Approved 10 September 2007.

EPCglobal: EPC Information Services (EPCIS) Version 1.0.1 Specification Approved September 21, 2007.

EPCglobal: Reader Management Standard 1.0.1, 3 May 31, 2007.

"EU data protection rules," European Commission, 2020. [Online]. Available: https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules_en. [Accessed 27 August 2020].

"EIP-1155: ERC-1155 Multi Token Standard," Ethereum Improvement Proposals, 2020. [Online]. Available: https://eips.ethereum.org/EIPS/eip-1155. [Accessed 10 September 2020].

E. J. Keogh, K. Chakrabarti, M. J. Pazzani, and S. Mehrotra. Dimensionality reduction for fast similarity search in large time series databases. Knowledge and Information Systems (KAIS), 3(3):263–286, 2001.

"EIP-20: ERC-20 Token Standard," 2020. [Online]. Available: https://eips.ethereum.org/EIPS/eip20. [Accessed 1 September 2020].

E. B. Association, "B2B Data Sharing: Digital Consent Management as a Driver for Data Opportunities," Euro Banking Association, 2018.

E. Commission, "European Commission," 2015. [Online]. Available: https://ec.europa.eu/info/law/payment-services-psd-2-directive-eu-2015-2366_en. [Accessed 10 November 2021].

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231.

ETSI GS CIM 009 V1.5.1 (2021-11) Context Information Management (CIM); NGSI-LD API https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.05.01_60/gs_CIM009v010501p.pdf.

Elliott, T. (2019). The State of the Octoverse: machine learning. GitHub. Retrieved 11 10, 2020, from https://github.blog/2019-01-24-the-state-of-the-octoverse-machine-learning/.

Edwards, L., & Veale, M. (2017). Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For. Duke Law & Technology Review, 16(18), 1–67. SSRN. dx.doi.org/10.2139/ssrn.2972855.

ECML-PKDD 2018 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, http://www.ecmlpkdd2018.org/wp-content/uploads/2018/09/567.pdf (http://www.ecmlpkdd2018.org/wp-content/uploads/2018/09/567.pdf).

«EC,» [Online]. Available: https://aws.amazon.com/ec2/instance-types/.

F. Vogelsteller and V. Buterin, "ERC-20 token standard," 2015. [Online]. Available: https://eips.ethereum.org/EIPS/eip-20 [Accessed 1-October-2020].

FAFT, Updated Guidance for a Risk-Based Approach to Virtual Assets and Virtual Asset Service Providers, https://www.fatf-gafi.org/publications/fatfrecommendations/documents/guidance-rba-virtual-assets-2021.html [accessed March 2022].

Feizi, Andisheh & Wong, Chui Yin. (2012). Usability of user interface styles for learning a graphical software application. 1089–1094. 10.1109/ICCISci.2012.6297188.

Final Report: Guidelines on the security measures for operational and security risks of payment services under Directive (EU) 2015/2366 (PSD2), European Banking Authority, Dec 2017.

Ferrari Dacrema, M., Cremonesi, P., & Jannach, D. (2019). Are we really making much progress? A worrying analysis of recent neural recommendation approaches. RecSys '19: Proceedings of the 13th ACM Conference on Recommender Systems, 101–109. ACM Digital Library. https://doi.org/10.1145/3298689.3347058.

Filice, S., Castellucci, G., Da San Martino, G., Moschitti, A., Croce, D., & Basili, R. (2018). KELP: a Kernel-based Learning Platform. Journal of Machine Learning Research 18(191):1–5.

Financial Stability Board (2017), Artificial intelligence and machine learning in financial services, www.fsb.org/wp-content/uploads/P011117.pdf.

Fine, S., Singer, Y., & Tishby, N. (1998). The hierarchical hidden Markov model: Analysis and applications. Machine Learning, 32, 41–62.

Fionda, V., & Pirró, G. (2019). triple2Vec: Learning Triple Embeddings from Knowledge Graphs. arxiv Preprint. arXiv. https://arxiv.org/abs/1905.11691.

Fortuna, B., Rupnik, J., Brank, J., Fortuna, C., Jovanoski, V., Karlovcec, M., Kazic, B., Kenda, K., Leban, G., Muhic, A., Novak, B., Novlian, J., Papler, M., Rei, L., Sovdat, B., Stopar, L., Grobelnik, M., & Mladenic, D. (2014). QMiner: Data Analytics Platform for Processing Streams of Structured and Unstructured Data. In Proceedings of the Software Engineering for Machine Learning Workshop at Neural Information Processing Systems (NIPS 2014).

Fasnacht, Daniel. (2018) "Open innovation ecosystems." Open Innovation Ecosystems. Springer, Cham, 2018. 131–172.

Fasnacht, Daniel. (2020) "The Ecosystem Strategy: Disruptive Business Model Innovation." Zeitschrift Führung und Organisation (zfo) 89.3: 168–173.

Florea, I. C., Vochin, O. A., Ciachir, L., & NAGEL-PICIORUS, C. P. (2017). Competition for Innovation in the Financial Software Industry-A Research on Hackathons. In BASIG INTERNATIONAL CONFERENCE: NEW TRENDS IN SUSTAINABLE BUSINESS AND CONSUMPTION (Vol. 2017, pp. 234–242).

Feldmann, A., & Teuteberg, F. (2020). Success factors for hackathons: German banks collaborate to tame the economic crisis. Journal of Business Strategy.

Flores, M. et al. (2018). How Can Hackathons Accelerate Corporate Innovation? In: Moon, I. et al. (eds.) Advances in Production Management Systems. Production Management for Data Driven, Intelligent, Collaborative, and Sustainable Manufacturing. pp. 167–175 Springer International Publishing, Cham.

Fasnacht, Daniel. (2019). "Open Innovation Ecosystem: The Winner Takes It All." In: Open Innovation Ecosystems. Management for Professionals. Springer, Cham. DOI 10 (2019): 978-3.

FINSEC Reference Architecture II – Available at https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ce3a941d&appId=PPGMS.

FBK, " D5.9 – Library of ML/DL Algorithms – III," p. 12.

FATF. (2019). Guidance for a risk-based approach, virtual assets and virtual asset service providers, FATF, 2019, Paris. https://www.fatf-gafi.org/publications/fatfrecommendations/documents/guidance-rba-virtual-assets.html.

"FIWARE," [Online]. Available: https://www.fiware.org/developers/.

F. Montesi and J. Weber, "Circuit breakers, discovery, and API gateways in microservices," 2016.

"FIWARE Vehicle Data model," [Online]. Available: https://fiwaredatamodels.readthedocs.io/en/latest/Transportation/Vehicle/Vehicle/doc/spec/index.html.

"FIWARE Alert Data Model," [Online]. Available: https://fiwaredatamodels.readthedocs.io/en/latest/Alert/doc/spec/index.html.

"FIWARE Weather Observed Data model," [Online]. Available: https://fiwaredatamodels.readthedocs.io/en/latest/Weather/WeatherObserved/doc/spec/index.html.

F. Khan, N. Akhtar and M. A. Qadeer, "RFID Enhancement in Road Traffic Analysis by Augmenting Reciever with TelegraphCQ," 2009 Second International Workshop on Knowledge Discovery and Data Mining, Moscow, 2009, pp. 331–334.

F. Moutinho, L. Paiva, J. Köpke, and P. Maló, "Extended Semantic Annotations for Generating Translators in the Arrowhead Framework," IEEE Transactions on Industrial Informatics, vol. 14, no. 6, pp. 2760–2769, Jun. 2018, doi: 10.1109/TII.2017.2780887.

F. V. a. V. Buterin, "ERC-20 token standard," 2015. [Online]. Available: https://eips.ethereum.org/EIPS/eip-20. [Accessed 19 November 2021].

Farrugia, S., Ellul, J., & Azzopardi, G. (2020). Detection of illicit accounts over the Ethereum blockchain. Expert systems with applications 150: 113318.

FATF. Guidance for a risk based approach, virtual assets and virtual asset service providers, FATF, paris. https://www.fatf-gafi.org/publications/fatfrecommendations/documents/guidance-rba-virtual-assets.html, 2019.

Frisendal, T. (2012). Design thinking for business analysis. In Design Thinking Business Analysis (pp. 15–24). Springer, Berlin, Heidelberg.

FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models (Accessed: January 2022).

FIWARE Foundation, "Smart Data Models," https://smartdatamodels.org/ (Accessed: January 2022).

FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Transportation (Accessed: January 2022).

FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Transportation/tree/master/Vehicle (Accessed: January 2022).
FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Weather/tree/master/WeatherObserved (Accessed: January 2022).
FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Transportation/tree/master/Road (Accessed: January 2022).
FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Transportation/tree/master/RoadSegment (Accessed: January 2022).
FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Alert/tree/master/Alert (Accessed: January 2022).
FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models, 2021. https://github.com/smart-data-models/dataModel.Transportation/tree/master/RoadAccident (Accessed: January 2022).
FIWARE Foundation, "Smart Data Models," https://github.com/smart-data-models/dataModel.Transportation/tree/master/RestrictedTrafficArea (Accessed: January 2022).
"FlinkCEP.", https://ci.apache.org/projects/flink/flink-docs-release-1.11/dev/libs/cep.html.
Gareth, J., Witten, D., Hastie, T., & Robert Tibshirani. (2013). An Introduction to Statistical Learning. Springer-Verlag New York. 10.1007/978-1-4614-7138-7.
Goldberger, J., Hinton, G., Roweis, S., & Salakhutdinov, R. R. (2004). Neighbourhood Components Analysis. Advances in neural information processing systems, 17(1), 513–520. https://proceedings.neurips.cc/paper/2004/file/42fe880812925e520249e808937738d2-Paper.pdf.
GDPR – IT Governance, https://www.itgovernance.co.uk/data-privacy/gdpr-overview/gdpr-faq/gdpr-scope (accessed Nov 2020).
GDPR scope – IT Governance, https://www.itgovernance.co.uk/data-privacy/gdpr-overview/gdprfaq/gdpr-scope (accessed Nov 2020).
G. Wondracek, T. Holz, E. Kirda, and C. Kruegel, "A Practical Attack to De-anonymize Social Network Users," in 2010 IEEE Symposium on Security and Privacy, May 2010, pp. 223–238.
G. Zhong and U. Hengartner, "A distributed k-anonymity protocol for location privacy," 2009 IEEE International Conference on Pervasive Computing and Communications. 2009, doi: 10.1109/percom.2009.4912774.
G. Wondracek, T. Holz, E. Kirda, and C. Kruegel, "A Practical Attack to De-anonymize Social Network Users," in 2010 IEEE Symposium on Security and Privacy, May 2010, pp. 223–238.
Growth Share Matrix https://fourweekmba.com/bcg-matrix/.

Gartner, "DevOps Gartener Glossary," Gartner, 2020. [Online]. Available: https://www.gartner.com/en/information-technology/glossary/devops.

Git, "Git," Software Freedom Conservancy, 2020. [Online]. Available: https://git-scm.com/about.

GitLab, "Gitlab docs," GitLab, 2020. [Online]. Available: https://docs.gitlab.com/charts/installation/deployment.html.

Google, "MLOps: Continuous delivery and automation pipelines in machine learning," Google, 2019. [Online]. Available: https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning.

Google, "Kubernetes-Kubectl," [Online]. Available: https://kubernetes.io/docs/tasks/tools/install-kubectl/.

Google, "Containers," Google, 2018. [Online]. Available: https://cloud.google.com/containers.

G. A. Stevens, B. Magri, D. Venturi and E. Andrade, "Redactable Blockchain – or – Rewriting History in Bitcoin and Friends," 2017 IEEE European Symposium on Security and Privacy (EuroS&P), pp. 111–126, 2017.

GDPR Recital (1).

Gaur, N., Desrosiers, L., Ramakrishna, V., Novotny, P., Baset, S.A., and O'Dowd, A. (2018). Hands-on Blockchain with Hyperledger: Building decentralized applications with Hyperledger Fabric and Composer. Packt Publishing.

GitLab homepage, GitLab, [Online]. Available: https://about.gitlab.com/.

Guidelines FIWARE data models https://fiware-datamodels.readthedocs.io/en/latest/guidelines/index.html.

GENE Ontology – bioinformatics initiative Accessible here: http://www.geneontology.org.

G. D. Orio, P. Maló, J. Barata, M. Albano, and L. L. Ferreira, "Towards a Framework for Interoperable and Interconnected CPS-populated Systems for Proactive Maintenance," in 2018 IEEE 16th.

G. Wood, "Ethereum: A Secure Decentralised Generalised Transaction Ledger, Ethereum Project Yellow Paper", 2018. [Online] Available at: https://ethereum.github.io/yellowpaper/paper.pdf. [Accessed: 20- October-2020].

Grafana, [Online]. Available: https://grafana.com/.

H. Hacigümüs, J. Sankaranarayanan, J. Tatemura, J. LeFevre, N. Polyzotis, "Odyssey: a multistore system for evolutionary analytics", PVLDB, vol. 6, pp. 1180–1181 (2013).

H. S. Pinto, S. Staab, and C. Tempich, "DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely-controlled and evolvInG Engineering of oNTologies," in Proceedings of the 16th European Conference on Artificial Intelligence, 2004, pp. 393–397.

Heitmann, B., Kinsella, S., Hayes, C. and Decker, S. Implementing Semantic Web Applications: Reference Architecture and Challenges. In International Workshop on Semantic Web enabled Software Engineering, collocated with the 8th International Semantic Web Conference (ISWC2009), 2009.

HELM, "Quickstart," HELM, 2020. [Online]. Available: https://helm.sh/docs/intro/quickstart/.

Henson, C. A., Pschorr, J. K., Sheth, A. P. and Thirunarayan, K. SemSOS: Semantic sensor observation service. Collaborative Technologies and Systems, International Symposium on, 0:44–53, 2009.

H. van der Veer and A. Wiles, "Achieving technical interoperability," European Telecommunications Standards Institute, 2008.

He, H. (2019, 10 10). The State of Machine Learning Frameworks in 2019. The Gradient. https://thegradient.pub/state-of-ml-frameworks-2019-pytorch-dominates-research-tensorflow-dominates-industry/.

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T.-S. (2017). Neural Collaborative Filtering. WWW '17: Proceedings of the 26th International Conference on World Wide Web, 173–182. ACM Digital Library. https://doi.org/10.1145/3038912.3052569.

Ho, T. K. (1995). Random decision forests (Proceedings of 3rd International Conference on Document Analysis and Recognition ed.). IEEE. doi.org/10.1109/ICDAR.1995.598994.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8). ACM Digital Library. doi.org/10.1162/neco.1997.9.8.1735.

Holton, G. A. (2014). Value-at-Risk: Theory and Practice (2nd ed.). London: Academic Press.

Hardjono, T., and Pentland, A. (2017). Open algorithms for identity federation. arXiv:1705.10880.

Hardjono, T., and Pentland, A. (2019). Data Cooperatives: Towards a foundation for decentralized personal data management. arXiv: 1905.08819.

Hu, V., Ferraiolo, D., Kuhn, D., Friedman, A., Lang, A., Cogdell, M.M., Schnitzer, A., Sandlin, K., Miller, R., and Scarfone, K. (2014). Attribute-Based Access Control Guide to Attribute Based Access Control (ABAC) Definition and Considerations. National Institute of Standards and Technology, NIST SP 800-162.

H. Xu, M. Chen, Y. Zhou, B. Du and L. Pan, "A Novel Comprehensive Quality Index QoX and the Corresponding Context-aware System Framework," 2018 IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China, 2018, pp. 2415–2419.

Hu, V., Kuhn, D., and Ferraiolo, D. (2015). Attribute-based access control. Comput. J. 48, 864–866. doi: 10.1109/MC.2015.33.

Harvard Business Review – You need an innovation strategy https://www.homeworkforyou.com/static_media/uploadedfiles/You%20Need%20an%20Innovation%20Strategy.pdf.

Heikki Topi and Allen Tucker. (2014). Computing Handbook, Third Edition: Information Systems and Information Technology. CRC Press.

High-Level Expert Group on Artificial Intelligence. (2018, June 14). European Commission. Retrieved January 2022, from https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligenc.

Huxley – The most in-demand skills for Fintech https://www.huxley.com/en-ae/blog/2020/04/the-most-in-demand-skills-for-fintech-covid/.

"Hyperledger Fabric – Hyperledger," 2020. [Online]. Available: https://www.hyperledger.org/use/fabric. [Accessed 5 September 2020].

Hong, Song-You, Yign Noh, and Jimy Dudhia. "A new vertical diffusion package with an explicit treatment of entrainment processes." Monthly weather review 134.9 (2006): 2318–2341.

H. Surendra and H. Mohan, "A review of synthetic data generation methods for privacy preserving data publishing," International Journal of Scientific & Technology Research, vol. 6, no. 3, pp. 95–101, 2017.

HAproxy, "HAproxy," HAproxy, [Online]. Available: http://www.haproxy.org/#desc.

Hashicorp, "Introduction to Terraform," Hashicorp, [Online]. Available: terraform.io/intro/index.html.

HELM, [Online]. Available: https://helm.sh/.

HCL, [Online]. Available: https://www.terraform.io/docs/configuration-0-11/syntax.html.

Harbor homepage, Linux Foundation, [Online]. Available: https://goharbor.io/.

Hubble, [Online]. Available: https://docs.cilium.io/en/v1.9/intro/#intro.

https://www.redflagalert.com/articles/data/what-is-the-fourth-aml-directive-aml4d (accessed in December 2020).

https://www.loc.gov/law/help/artificial-intelligence/regulation-artificial-intelligence.pdf (accessed in December 2020).

https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework (accessed in December 2020).

https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf (accessed in December 2020).

https://docs.docker.com/get-started/overview/, Docker. [Online]. [Accessed 24 03 2021].

https://www.oreilly.com/library/view/software-architecture-with/9781786468529/ch08s04.html.
http://www.dossier-andreas.net/software_architecture/index.html.
http://www.se.rit.edu/~emad/teaching/slides/CISC322_06_ArchitectureStyles_sep20.pdf.
https://sarasanalytics.com/blog/data-pipeline-architecture.
https://www.astera.com/type/blog/data-pipeline-architecture.
https://airflow.apache.org/docs/apache-airflow/stable/index.html.
https://luigi.readthedocs.io/en/stable.
https://www.knime.com/.
https://c3.ai/products/c3-ai-ex-machina/.
https://pipeline.loni.usc.edu/learn/user-guide/building-a-workflow/.
https://streampipes.apache.org/.
https://advisera.com/27001academy/what-is-iso-27001/, page 1 (accessed in December 2020).
https://fcicyber.com/top-5-ways-the-financial-services-industry-can-leverage-nist-for-cybersecuritycompliance/ (accessed in December 2020).
https://financialservices.mazars.com/gdpr-psd2-issues-fintechs/ (accessed in December 2020).
https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/digitalbanking/psd2 (accessed in December 2020).
https://eba.europa.eu/sites/default/documents/files/documents/10180/2060117/d53bf08f-990b-47ba-b36f-15c985064d47/Final%20report%20on%20EBA%20Guidelines%20on%20the%20security%20measures%20for%20operational%20and%20security%20risks%20under%20PSD2%20(EBA-GL-2017-17).pdf (accessed in December 2020).
https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/digitalbanking/psd2 (accessed in December 2020).
https://www.investopedia.com/terms/m/mifid-ii.asp (accessed in December 2020).
https://www.redflagalert.com/articles/data/what-is-the-fourth-aml-directive-aml4d (accessed in December 2020).
https://www.loc.gov/law/help/artificial-intelligence/regulation-artificial-intelligence.pdf (accessed in December 2020).
https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework (accessed in December 2020).
https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf (accessed in December 2020).
https://www.fiware.org/developers/smart-data-models/.
https://fiware-orion.readthedocs.io/en/master/.

https://fiware.github.io/specifications/ngsiv2/stable/.
https://quantumleap.readthedocs.io/en/latest/.
https://app.swaggerhub.com/apis/smartsdk/ngsi-tsdb/0.1.
https://konghq.com/kong/.
https://sumo.dlr.de/docs/index.html.
https://opendata.aemet.es/.
https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/.
https://soilgrids.org/#!/?layer=TAXNWRB_250m&vector=1.
https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs.
http://crowdpolicy.com/.
https://www.standardsplusinnovation.eu/, retrieved 12/12/2020.
https://databox.com/new-vs-returning-visitors.
https://ec.europa.eu/info/publications/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics-0_en.
https://www.eventbrite.com/blog/asset/ultimate-way-reduce-no-shows-free-events/.
https://ec.europa.eu/digital-single-market/en/news/ict-verticals-and-horizontals-blockchain-standardisatio.
https://marketinginsidergroup.com/content-marketing/why-social-media-is-important-for-business-marketing/.
https://www.webfx.com/internet-marketing/why-is-social-media-so-important.
https://www.campaignmonitor.com/resources/guides/why-email/.
https://www.campaignmonitor.com/blog/email-marketing/8-epic-email-newsletters-you-can-learn-from/.
https://v12data.com/blog/top-strategies-to-personalize-your-email-campaigns/&sa=D&source=editors&ust=1623772876236000&usg=AOvVaw3QVzfE7H7Qh5tpWHBVMOum.
https://thenextscoop.com/company-host-hackathon/.
https://technology.finra.org/articles/why-enterprises-should-host-hackathons.html.
https://www.livechat.com/success/company-can-benefit-from-hackathons/.
https://digitalagencynetwork.com/webinar-marketing-strategy-how-to-promote-your-upcoming-webinar/.
https://www.meraevents.com/blog/benefits-of-attending-workshops-for-professionals-businessowners.
https://yellowdoorcollective.com/blog/3-reasons-why-hosting-a-workshop-will-add-value-toyour-business/.
https://marketplace.infinitech-h2020.eu/workshops/stakeholders-workshops-series-1.

https://digital-strategy.ec.europa.eu/en/activities/edihs.
https://advisera.com/27001academy/what-is-iso-27001/, page 1 (accessed in December 2020).
https://fcicyber.com/top-5-ways-the-financial-services-industry-can-leverage-nist-for-cybersecurity-compliance/ (accessed in December 2020).
https://financialservices.mazars.com/gdpr-psd2-issues-fintechs/ (accessed in December 2020).
https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/digital-banking/psd2 (accessed in December 2020).
https://eba.europa.eu/sites/default/documents/files/documents/10180/2060117/d53bf08f-990b-47bab36f-15c985064d47/Final%20report%20on%20EBA%20Guidelines%20on%20the%20security%20measures%20for%20operational%20and%20security%20risks%20under%20PSD2%20(EBA-GL-2017-17).pdf (accessed in December 2020).
https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/digital-banking/psd2 (accessed in December 2020).
https://www.investopedia.com/terms/m/mifid-ii.asp (accessed in December 2020).
https://rancher.com/docs/rancher/v2.5/en/installation/requirements/#hardware-requirem.
https://www.abe-eba.eu/thought-leadership-innovation/open-banking-working-group/management-summary-artificial-intelligence-in-the-era-of-open-banking/.
International Organization for Standardization, "ISO/IEC 29115:2013 Information technology – Security techniques – Entity authentication assurance framework." 2013, [Online]. Available: https://www.iso.org/standard/45138.html.
IBM – Security, Policy and Compliance, https://www.ibm.com/cloud/garage/architectures/securityArchitecture/security-policy-governance-risk-compliance [accessed March 2022].
ICAO, "ICAO Document 9303 – Machine Readable Travel Documents." [Online]. Available: https://www.icao.int/publications/pages/publication.aspx?docnum=9303. [Accessed: 07-Jul-2021].
I.T. Gartner, "Glossary (2020),'Data governance,' " URL: https://www.gartner.com/en/information-technology/glossary/data-governance.
Innovation Project Portfolio Management: A Meta-Analysis, June 2012, International Journal of Product Development: 77–94. https://www.researchgate.net/publication/259557845_Innovation_Project_Portfolio_Management_A_Meta-Analysis.
Innovation management techniques and tools: a review from theory and practice, Antonio Hidalgo and Jose Albors Department of Business Administration.

Universidad Politecnica de Madrid, c/Jose Gutierrez https://oa.upm.es/2406/1/INVE_MEM_2008_55568.pdf.

IBM, "Containerization," [Online]. Available: https://www.ibm.com/cloud/learn/containerization. [Accessed 24 3 2021].

IBM, "Containers," Google, 2018. [Online]. Available: https://www.ibm.com/cloud/learn/containerization. [Accessed 24 03 2021].

Intellectual Property Rights https://www.researchgate.net/figure/Intellectual-PropertyReadiness-Level-Part-of-TRL-Hasenauer-et-al-Managing_fig4_313063121.

IBM Institute for Business Value, Moving to a token-driven economy Enabling the digitization of real-world assets, 2018. [Online] Available at: https://www.ibm.com/downloads/cas/YMRKPOJ8 [Accessed 15-October-2020].

IBM digital transformation https://www.ibm.com/thought-leadership/institute-business-value/report/digital-transformation#.

Istio, "What is Istio," Istio, 2020. [Online]. Available: https://istio.io/latest/docs/concepts/what-is-istio/.

International Conference on Industrial Informatics (INDIN), Jul. 2018, pp. 146–151, doi: 10.1109/INDIN.2018.8472041.

"Introduction — hyperledger-fabricdocs master documentation," Hyperledger, 2020. [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.2/whatis.html. [Accessed 03 September 2020].

I. Karagiannis, K. Mavrogiannis, J. Soldatos, D. Drakoulis, E. Troiano and A. Polyviou, "Blockchain Based Sharing of Security Information for Critical Infrastructures of the Finance Sector," Computer Security Lecture Notes in Computer Science, Computer Security. IOSEC 2019, MSTEC 2019, FIN-SEC 2019, vol. 11981, pp. 226–241, 2020.

Iacono, Michael J., et al. "Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models." Journal of Geophysical Research: Atmospheres113.D13 (2008).

IEEE. (2019). ETHICALLY ALIGNED DESIGN A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems First Edition. Retrieved January 2022, from https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf.

IEEE Blockchain Standards, https://blockchain.ieee.org/standards (Accessed: January 2022).

IEEE Blockchain Standards, https://blockchain.ieee.org/standards (Accessed: January 2022).

IEEE Big Data Governance and Metadata Management (BDGMM) https://standards.ieee.org/industry-connections/BDGMM-index.html Accessed: January 2022).

I. Cambridge Semantics, "Leveraging Semantic and Graph Technology to Tame the Enterprise Data Storm." combridgesemantics.com, Jun. 2019, Available Online: https://blog.cambridgesemantics.com/.

International Organization for Standardization. https://www.iso.org/home.html (Accessed: January 2022).

ISO – ISO/IEC JTC 1 – Information Technology. ISO. Retrieved January, 2022 from https://www.iso.org/isoiec-jtc-1.html.

Information security, cybersecurity and privacy protection. JTC 1. Retrieved January, 2022 from https://jtc1info.org/technology/subcommittees/information-security-cybersecurity-privacy-protection/.

Imerman, M. B., & Fabozzi, F. J. (2020). A Conceptual Framework for FinTech Innovation. Available at SSRN 3543810.

ISO/IEC Guide 59:2019(en): ISO and IEC recommended practices for standardization by national bodies.

i-SCOOP – Digital transformation strategy: the bridges to build https://www.i-scoop.eu/digitaltransformation/digital-transformation-strategy/.

IETF – Uniform Resource Names Accessible here: http://tools.ietf.org/html/rfc2141 [W3C-RDF] World Wide Web Consortium – Resource Description Framework, Accessible here: http://www.w3.org/TR/rdf-syntax-grammar/.

J. Dumortier, "REGULATION (EU) NO 910/2014 ON ELECTRONIC IDENTIFICATION AND TRUST SERVICES FOR ELECTRONIC TRANSACTIONS IN THE INTERNAL MARKET (EIDAS REGULATION)," in EU Regulation of E-Commerce, Edward Elgar Publishing, 2017.

J. Froehlich and J. Krumm, "Route prediction from trip observations, Soc," Automot. Eng. Spec. Publ, vol. 2193, p. 53, 2008.

J. Krumm and E. Horvitz, "Predestination: Inferring Destinations from Partial Trajectories," in UbiComp 2006: Ubiquitous Computing, 2006, pp. 243–260.

J. Duggan, A. J. Elmore, M. Stonebraker, M. Balazinska, B. Howe, J. Kepner, S. Madden, D. Maier, T. Mattson, S. Zdonik, "The BigDAWG polystore system", SIGMOD Record, vol. 44, no. 2, pp. 11–16, 2015.

J. LeFevre, J. Sankaranarayanan, H. Hacıgümüs, J. Tatemura, N. Polyzotis, M. Carey, "MISO: souping up big data query processing with a multistore system", in ACM SIGMOD, pp. 1591–1602 (2014).

J. Wang, T. Baker, M. Balazinska, D. Halperin, B. Haynes, B. Howe, D. Hutchison, S. Jain, R. Maas, P. Mehta, D. Moritz, B. Myers, J. Ortiz, D. Suciu, A. Whitaker, S. Xu, "The Myria big data management and analytics system and cloud service", in Conference on Innovative Data Systems Research (CIDR) (2017).

J. Krumm, "Inference Attacks on Location Tracks," in Pervasive Computing, 2007, pp. 127–143.

J. Köpke, "Annotation paths for matching xml-schemas," Data & Knowledge Engineering, vol. 122, pp. 25–54, 2019.

J. Köpke and J. Eder, "Semantic annotation of XML-schema for document transformations," in OTM Confederated International Conferences "On the Move to Meaningful Internet Systems", 2010, pp. 219–228.

J. Dafflon, J. Baylina, and T. Shababi, "ERC-777 token standard," 2015. [Online]. Available: https://eips.ethereum.org/EIPS/eip-777 [Accessed 10-October-2020].

Jolliffe, I.T. (2002). Principal Component Analysis, second edition, New York: Springer-Verlag New York, Inc.

Jiménez, Pedro A., et al. "A revised scheme for the WRF surface layer formulation." Monthly Weather Review 140.3 (2012): 898–918.

«Jenkins homepage,» [Online]. Available: https://www.jenkins.io/.

J. Wang, Y. Luo, Y. Zhao, and J. Le, "A Survey on Privacy Preserving Data Mining," in 2009 First International Workshop on Database Technology and Applications, Apr. 2009, pp. 111–114.

Jacobs, I. and Walsh, N. Architecture of the World Wide Web, Volume One, World Wide Web Consortium, Recommendation REC-webarch-20041215, 2004.

J.Y. Lee, "A decentralized token economy: How blockchain and cryptocurrency can revolutionize business", Business Horizons (2019). [Online]. Available at: www.sciencedirect.com [Accessed: 10- October-2020].

Jenkins, "Jenkins," [Online]. Available: https://jenkins.io/.

Joining forces for blockchain standardisation https://europa.eu/newsroom/events/joining-forces-blockchain-standardisation_en (Accessed: January 2022).

James Moar: "Three Trends Accelerating the Growth of Digital Identity", Juniper Research, July 2019.

Jeff Galvin et al., "Synergy and disruption: Ten trends shaping Fintech", McKinsey & CO, December 2018.

Keycloak – Authorization Services Guide, https://www.keycloak.org/docs/latest/authorization_services/#authorization-services, [Accessed: March 2022].

KIE API, https://docs.jboss.org/drools/release/6.3.0.Final/kie-api-javadoc/overview-summary.html, [accessed March 2022].

K. K. Petros Kavassalis, "D4.3 Operational and Technical Documentation of SP (ATHEX, Hellenic Post) integration (production)." 2018, [Online]. Available: http://www.lepsproject.eu/sites/default/files/leps/public/contentfiles/deliverables/LEPS%20D4.3%20Operational%20and%20Technical%20Documentation%20of%20SP%20integration.pdf.

K. Mivule, "Utilizing Noise Addition for Data Privacy, an Overview," arXiv [cs.CR], Sep. 16, 2013.

KPMG – Insurtech 10: Trends for 2019 https://assets.kpmg/content/dam/kpmg/xx/pdf/2019/03/insurtech-trends-2019.pdf.

K. Mivule, "Utilizing Noise Addition for Data Privacy, an Overview," arXiv [cs.CR], Sep. 16, 2013.

Kubernetes, "What is kubernetes?," Kubernetes, [Online]. Available: https://kubernetes.io/docs/concepts/overview/what-is-kubernetes/. [Accessed 24 03 2021].

K. Philippe, "Architectural Blueprints – The "4+1" View Model of Software Architecture", IEEE Software, vol. 12, pp. 42–50, November 1995.

Kubernetes, "What is kubernetes?," Kubernetes, 5 August 2020. [Online]. Available: https://kubernetes.io/docs/concepts/overview/what-is-kubernetes/.

Kubernetes – Namespaces, [Online]. Available: https://kubernetes.io/docs/concepts/overview/workingwith-objects/namespaces/.

kfp kubeflow sdk, [Online]. Available: https://www.kubeflow.org/docs/components/pipelines/sdk/sdk-overview/.

kfp.dsl package, [Online]. Available: https://kubeflowpipelines.readthedocs.io/en/latest/source/kfp.dsl.html.

K. K. Framework. [Online]. Available: https://www.kubeflow.org/docs/external-add-ons/kserve/kserve/.

Kubernetes, [Online]. Available: https://kubernetes.io/. [Accessed 31 05 2021].

Kubeflow, "About Kubeflow," Kubeflow, 2020. [Online]. Available: https://www.kubeflow.org/docs/about/kubeflow/.

K. Awada, M. Eltabakh, C. Tang, M. Al-Kateb, S. Nair, G. Au, "Cost Estimation Across Heterogeneous SQL-Based Big Data Infrastructures in Teradata IntelliSphere", in EDBT, pp. 534–545 (2020).

Kılıç, B., Özturan, C., & Şen, A. (2021). Finding convex subgraphs in Blockchain transaction graphs, submitted to BCCA 2021 conference (under review).

K. I. P. L. D. V. C. J. P. K. A. Bhaskaran, "Double-blind consent-driven data sharing on blockchain.," in 2018 IEEE International Conference on Cloud Engineering (IC2E), 2018.

Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R.G.L., El Rouayheb, S., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P.B., Gruteser, M., Harchaoui, Z., He, C., He, L., Huo, Z., Hutchinson, B., Hsu, J., Jaggi, M., Javidi, T., Joshi, G., Khodak, M., Konecný, J., Korolova, A, Koushanfar, F., Koyejo, S., Lepoint, T., Liu, Y., Mittal, P., Mohri, M., Nock, R., Özgür, A., Pagh, R., Raykova, M., Qi, H., Ramage, D., Raskar, R., Song, D., Song, W., Stich, S.U., Sun, Z., Suresh, A.T., Tramér, F., Vepakomma, P., Wang, J., Xiong, L., Xu, Z., Yang, Q., Yu, F.X., Yu, H.,

and Zhao, S. (2019). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. arXiv:1912.04977.

K. Tuononen, "The impact of PSD2 directive on the financial services industry.," 2019.

KPMG, "KPMG," 2021. [Online]. Available: https://home.kpmg/ph/en/home/insights/2019/07/openbanking-opens-opportunities-for-greater-value.html. [Accessed 15 November 2021].

Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R.G.L., El Rouayheb, S., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P.B., Gruteser, M., Harchaoui, Z., He, C., He, L., Huo, Z., Hutchinson, B., Hsu, J., Jaggi, M., Javidi, T., Joshi, G., Khodak, M., Konecný, J., Korolova, A, Koushanfar, F., Koyejo, S., Lepoint, T., Liu, Y., Mittal, P., Mohri, M., Nock, R., Özgür, A., Pagh, R., Raykova, M., Qi, H., Ramage, D., Raskar, R., Song, D., Song, W., Stich, S.U., Sun, Z., Suresh, A.T., Tramér, F., Vepakomma, P., Wang, J., Xiong, L., Xu, Z., Yang, Q., Yu, F.X., Yu, H., and Zhao, S. (2021). Advances and open problems in federated learning. Foundations and trends in machine learning, 14 (1–2), pp. 1–210.

KPMG, "Open banking opens opportunities for greater customer," KPMG, 2021.

Kruchten, Philippe (1995, November). Architectural Blueprints – The "4+1" View Model of Software Architecture. IEEE Software 12 (6), pp. 42–50.

Komulainen, H., Makkonen, H. Customer experience in omni-channel banking services. J Financ Serv Mark 23, 190–199 (2018). https://doi.org/10.1057/s41264-018-0057-6.

Kumar, N., Singh, A., Handa, A., & Shukla, S.K. (2020). Detecting malicious accounts on the Ethereum blockchain with supervised learning. In Proceedings of International Symposium on Cyber Security Cryptography and Machine Learning, pp. 94–109. Springer, Cham.

K. Chan and A. W. Fu. E cient time series matching by wavelets. In Proceedings of the International Conference on Data Engineering (ICDE), pages 126–133. IEEE Computer Society, 1999.

K. W. Ong, Y. Papakonstantinou, and R. Vernoux, "The SQL++ semi-structured data model and query language: a capabilities survey of SQL-on-Hadoop, NoSQL and NewSQL databases", CoRR, abs/1405.3631, 2014.

Kubernetes IP management https://kubernetes.io/docs/concepts/services-networking/service/#choosing-your-own-ip-address.

Kubernetes. [Online] https://kubernetes.io.

Kubernetes IP management [Online]. Available: https://kubernetes.io/docs/concepts/servicesnetworking/service/#choosing-your-own-ip-address.

Konstantinos Chatzikokolakis, Ehab Elsalamouny, Catuscia Palamidessi. Efficient Utility Improve-ment for Location Privacy. Proceedings on Privacy Enhancing Technologies, De Gruyter Open, 2017, 2017(4), pp. 308–328. 10.1515/popets-2017-0051.hal-01422842v2.

Kathleen Brandenburg. Design as Business Strategy: Key Principles of Successful Design Thinking https://iacollaborative.com/perspective/design-as-business-strategy-key-principles-of-successful-design-thinking/.

L. Haas, D. Kossmann, E. Wimmers, J. Yang. Optimizing Queries across Diverse Data Sources. Int. Conf. on Very Large Databases (VLDB), pp. 276–285 (1997).

Lundberg, S.M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), pp. 4768–4777.

L. Neumeyer, B. Robbins, A. Nair, and A. Kesari, "S4: Distributed Stream Computing Platform," in 2010 IEEE International Conference on Data Mining Workshops, 2010, pp. 170–177.

Longerstaey, J., & Spencer, M. (1996). Riskmetrics — technical document. Morgan Guaranty Trust Company of New York: 51, 54.

LeewayHertz – How can blockchain simplify KYC and AML processes? https://www.leewayhertz.com/blockchain-in-aml/.

L. Encrypt, "Let's Encrypt," [Online]. Available: https://letsencrypt.org/.

Le-Phuoc, D., Dao-Tran, M. Parreira, J. X. and Hauswirth, M. A Native and Adaptive Approach for Unified Processing of Linked Streams and Linked Data. Proceedings of the 10th International Conference on The Semantic Web (ISWC'11), Springer, 2011.

Le-Phuoc, D., Nguyen Mau, H., Parreira, J. X. and Hauswirth, M.. The Linked Sensor Middleware – Connecting the Real World and the Semantic Web. Proceedings of the 10th International Conference on The Semantic Web (ISWC'11), Springer, 2011.

Le-Phuoc, D. and Hauswirth, M. Linked open data in sensor data mashups. Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09) in conjunction with ISWC 2009.

Liu, F.T., Ting, K.M., & Zhou, Z.-H. (2008). Isolation Forest. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, pp. 413–422.

Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y. P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., & Wiessner, E. (2018). Microscopic Traffic Simulation using SUMO. 21st International Conference on Intelligent Transportation Systems (ITSC), 2575–2582. IEEE Xplore. /doi.org/10.1109/ITSC.2018.8569938.

L. Oliveira, L. Zavolokina, I. Bauer, and G. Schwabe, "To Token or not to Token: Tools for Understanding Blockchain Tokens", Thirty Ninth International Conference on Information Systems, San Francisco 2018.

L. Monso, "Asset tokenization: What is it and why does it matter?", 2019. [Online]. Available at: https://medium.com/@Metaco/asset-tokenization-what-is-it-and-why-does-it-matter-3f6892273dfe [Accessed: 20-October-2020].

L. Monso, "Asset tokenization: Benefits and challenges ahead", 2019 [Online]. Available at: https://medium.com/@Metaco/asset-tokenization-benefits-and-challenges-ahead-475d737eba57 [Accessed: 20-October-2020].

Long G., Tan Y., Jiang J., Zhang C. (2020) Federated Learning for Open Banking. In: Yang Q., Fan L., Yu H. (eds) Federated Learning. Lecture Notes in Computer Science, vol 12500. Springer, Cham. https://doi.org/10.1007/978-3-030-63076-8_17.

Landgren, P., Srivastava, V., and Leonard, N.E. (2016). On distributed cooperative decisionmaking in multi armed bandits. In Proceedings of the IEEE European Control Conference (ECC), 243–248.

"Leveraging Semantic and Graph Technology to Tame the Enterprise Data Storm." https://info.cambridgesemantics.com/hubfs/whitepapers/Perfect%20Storm%20Fin%20Svcs%20Whitepaper.pdf (accessed Mar. 26, 2022).

Laudet, D., (2017). Infographic: Worldwide Hackathon Figures in 2016, Available at: http://agency.bemyapp.com/insights/infographics-hackathon-figures-in-2016.html.

Lafourcade, P., & Lombard-Platet, M. (2020). About blockchain interoperability. Information Processing Letters, 161, 105976.

Links, Skip. "High-Resolution Land Use and Land Cover Mapping".

M. Armbrust, R. Xin, C. Lian, Y. Huai, D. Liu, J. Bradley, X. Meng, T. Kaftan, M. Frank-lin, A. Ghodsi, M. Zaharia, "Spark SQL: relational data processing in Spark", in ACM SIGMOD, 2015, pp. 1383–1394.

Manisha Patel, "Top Five Impacts of GDPR on Financial Services" , Fintech Times, November 2017.

Mark Halstead, "What is the Fourth AML Directive (AML4D)-RedFlagAlert", https://www.redflagalert.com/articles/data/what-is-the-fourth-aml-directive-aml4d.

Martin David, Jorge Bernal, Julien Bringer, Nicolas Notario, Eduardo Gonzales – D3.1 – ARIES eID ecosystem technical design – July 2017.

Marcus Brandenburger, Eduarda Freire, "Witdom Project, D4.2 – Final specification of an end-to-end secure architecture", August 2016. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ac7baf1a&appId=PPGMS [Accessed 10-July-2021].

M. Gruteser and B. Hoh, "On the Anonymity of Periodic Location Samples," in Security in Pervasive Computing, 2005, pp. 179–192.

M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability," Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security – CCS '13. 2013, doi: 10.1145/2508859.2516735.

M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability," Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security – CCS '13. 2013, doi: 10.1145/2508859.2516735.

M. Duckham and L. Kulik, "A Formal Model of Obfuscation and Negotiation for Location Privacy," Lecture Notes in Computer Science. pp. 152–170, 2005, doi: 10.1007/11428572_10.

M. Duckham and L. Kulik, "Location privacy and location-aware computing," Dynamic & mobile GIS: investigating change in, 2006, [Online]. Available: https://www.academia.edu/download/44974954/Location_privacy_and_locationaware_comp20160421-7519-143glj.pdf.

Miguel E. Andrés and Nicolás Emilio Bordenabe and Konstantinos Chatzikokolakis and Catuscia Palamidessi (2012). Geo-Indistinguishability: Differential Privacy for Location-Based Systems. CoRR, abs/1212.1984.

M. Gruteser and D. Grunwald, "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking," Proceedings of the 1st international conference on Mobile systems, applications and services – MobiSys '03. 2003, doi: 10.1145/1066116.1189037.

M. Hirzel et al., "IBM Streams Processing Language: Analyzing Big Data in motion," in IBM Journal of Research and Development, vol. 57, no. 3/4, pp. 7:1–7:11, May-July 2013.

Miklos Dietz et al., "Cutting through the noise around financial technology", McKinsey & Co, February 2016.

Michael Lamer: "The Future of FINTECH ∼ The new Standard", Juniper Research, May 2019.

More Than Digital – Innovation – Definition, Innovation Types And Meaning https://morethandigital.info/en/innovation-definition-innovation-types-and-meaning/.

MIT Sloan Management Review https://sloanreview.mit.edu/projects/strategy-drives-digital-transformation/.

Meng, Z., McCreadie, R., Macdonald, C., Ounis, I., Liu, S., Wu, Y.,… & Zhang, Q. (2020, September). Beta-rec: Build, evaluate and tune automated recommender systems. In Fourteenth ACM conference on recommender systems (pp. 588–590).

"Microservices a definition of this new architectural term," [Online]. Available: https://martinfowler.com/articles/microservices.html#footnote-etymology. [Accessed 24 03 2021].

"Microservices.io," [Online]. Available: https://microservices.io/. [Accessed 20 October 2020].

M. Dudjak and G. Martinović, "An API-first methodology for designing a microservice-based Backend as a Service platform.," Information Technology and Control, vol. 49, no. 2, pp. 206–223, 2020.

M. J. T. D.-M. J. B. S. A. M. N. T. M. I. … & D.-L.-H.-V. E. Naeem, "Trends and Future Perspective Challenges in Big Data. In Advances in Intelligent Data Analysis and Applications," in Springer, Singapore, 2021.

Medium – How to: Business Model Canvas explained https://medium.com/seed-digital/how-to-business-model-canvas-explained-ad3676b6fe4a.

Medium – Introduction to Lean Canvas https://medium.com/@steve_mullen/an-introduction-to-lean-canvas-5c17c469d3e0.

M. Noura, M. Atiquzzaman, and M. Gaedke, "Interoperability in internet of things: Taxonomies and open challenges," Mobile Networks and Applications, vol. 24, no. 3, pp. 796–809, 2019.

M. Elkhodr, S. Shahrestani, and H. Cheung, "The Internet of Things: vision & challenges," in IEEE 2013 Tencon-Spring, 2013, pp. 218–222.

M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "Methontology: from ontological art towards ontological engineering," 1997.

M. Niranjanamurthy, B. N. Nithya and S. Jagannatha, "Analysis of Blockchain technology: pros, cons and SWOT," Cluster Computing, vol. 22, no. S6, pp. 14743–14757, 2018.

M. Casey, J. Crane, G. Gensler, S. Johnson, N. Narula and C. A. Wyplosz, The impact of blockchain technology on finance, Geneva: International Center for Monetary and Banking Studies (ICMB), 2018.

Muselli, M. (2006). Springer Berlin Heidelberg. In Neural Nets (pp. 23–30). Springer Berlin Heidelberg. doi.org/10.1007/11731177_4.

M. di Angelo and G. Salzer, "Tokens, Types, and Standards: Identification and Utilization in Ethereum" IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS), 2020.

Mun, M., Hao, S., Mishra, N., Shilton, K., Burke, J., Estrin, D., Hansen, M., and Govindan, R. (2010). Personal data vaults: A locus of control for personal data streams. In Proceedings of the 6th International Conference Co-NEXT '10, 1–12.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., Aguera y Arcas, B. (2017). CommunicationEfficient Learning of Deep Networks from Decentralized Data, AISTAT.

Miguel Lara and Kate Lockwood. (2016). Hackathons as Community-Based Learning: a Case Study. TechTrends 60, 5 (2016), 486–495.

M. Jordan. (2012) Planning a hackfest. Open Data Learning Summit.

Mueller, J. (2017). FinTech: Considerations on How to Enable a 21st Century Financial Services Ecosystem. Viewpoints], Milken Institute.

Mention, A. L. (2021). The age of FinTech: Implications for research, policy and practice. The Journal of FinTech, 1(01), 2050002.

mlflow.org, "An open source platform for the machine learning lifecycle," [Online]. Available: https://mlflow.org/.

M. Westerkamp, F. Victor, and A. Küpper, "Tracing manufacturing processes using blockchain-based token compositions", Digital Communications and Networks, 2019.

MetaCX for Healthcare. [Online]. Available: https://metacx.com/healthcare.

Marko Komssi, Danielle Pichlis, Mikko Raatikainen, Klas Kindström, and Janne Järvinen. (2015). What are Hackathons for? IEEE Software 32, 5 (2015), 60–67.

«MetalLB,» [Online]. Available: https://metallb.org/.

Mahendraratnam, N., Sorenson, C., Richardson, E., Daniel, G.W., Buelt, L., Westrich, K., Qian, J., Campbell, H., McClellan, M., and Dubois, R.W. (2019). Value-based arrangements may be more prevalent than assumed. The American Journal of Managed Care, 25(2), pp. 70–76.

N. Li, T. Li, and S. Venkatasubramanian, "t-Closeness: Privacy Beyond k-Anonymity and l-Diversity," in 2007 IEEE 23rd International Conference on Data Engineering, Apr. 2007, pp. 106–115.

Netguru – Top 7 Technology Trends in Fintech in 2021 https://www.netguru.com/blog/top-7-technology-trends-in-fintech-in-2021.

"NGSI specifications," [Online]. Available: http://fiware.github.io/specifications/ngsiv2/stable/.

N. Gaur, L. Desrosiers, V. Ramakrishna, P. Novotny, SA. Baset, and A O'Dowd, Hands-On Blockchain with Hyperledger, 2018.

NCEP. (2015). NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. U.S.: Weather Service/NOAA/U.S.

Niu, Guo-Yue, et al. "The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local scale measurements." Journal of Geophysical Research: Atmospheres 116.D12 (2011).

Next-gen Technology transformation in Financial Services https://www.mckinsey.com/~/media/mckinsey/industries/financial%20services/our%20insights/nex

t-gen%20technology%20transformation%20in%20financial%20services/next-gen-technology-transformation-in-financial-services.pdf.

N. Guarino, Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6–8, Trento, Italy, vol. 46. IOS press, 1998.

N. Kapsoulis, A. Psychas, G. Palaiokrassas, A. Marinakis, A. Litke and T. Varvarigou, "Know Your Customer (KYC) Implementation with Smart Contracts on a Privacy-Oriented Decentralized Architecture," Future Internet, vol. 12, no. 2, p. 41, 2020.

National Institute of Standards and technologies https://www.nist.gov/ (Accessed: January 2022).

Nick Taylor and Loraine Clarke. (2018). Everybody's Hacking: Participation and the Mainstreaming of Hackathons. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 172.

Nolte, A., Chounta, I. A., & Herbsleb, J. D. (2020). What Happens to All These Hackathon Projects? Identifying Factors to Promote Hackathon Project Continuation. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), 1–26.

Osterwalder, A. P. (2015). Value Proposition Design. New York: Wiley. Retrieved from Osterwalder, A., Pigneur, Y., Bernarda, G., Smith, A., & Papadakos, T. (2015). Value Proposition Design. New York: Wiley.

Osterwalder, A., & Pigneur, Y. (2013). Business Model Generation. Hoboken, NJ: Wiley.

Oehmichen, A., Jain, S., Gadotti, A., and de Montjoye, Y.-A. (2019). OPAL: High performance platform for large-scale privacy-preserving location data analytics. In Proceedings of BigData 2019: 1332–1342.

Osterwalder, A., Pigneur, Y. et al. (2010). 'Business Model Generation', self published.

Osterwalder, Alexander; Pigneur, Yves. (2013). Business Model Generation. Hoboken, NJ: Wiley.

OECD (2019), "Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)", OECD Digital Economy Papers, No. 291, OECD Publishing, Paris, https://doi.org/10.1787/d62f618a-en.

Oehmichen, A., Jain, S., Gadotti, A., and de Montjoye, Y.-A. (2019). OPAL: High performance platform for large-scale privacy-preserving location data analytics. In Proceedings of BigData 2019: pp. 1332–1342.

"Open API specification," 2020. [Online]. Available: https://swagger.io/specification/#:~:text=The%20OpenAPI%20Specification%20(OAS)%20defines,or

%20through%20network%20traffic%20inspection. [Accessed 20 October 2020].
"OData (Open Data Protocol)," [Online]. Available: https://www.odata.org/. [Accessed 24 03 2021].
"OpenAPI Specification," [Online]. Available: http://spec.openapis.org/oas/v3.0.3. [Accessed 24 03 2021].
[Online]. Available: Lead Beneficiary, Contributor, Internal Reviewer, Quality Assurance.
[Online]. Available: Can be left void.
[Online]. Available: https://en.wikipedia.org/wiki/Bastion_host.
OpenLDAP, "OpenLDAP," [Online]. Available: https://www.openldap.org/.
«Oralce Virtual Box,» [Online]. Available: https://www.virtualbox.org/.
«Open-LDAP,» Openldap Foundation, [Online]. Available: https://www.openldap.org/.
OECD (2019), "Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)", OECD Digital Economy Papers, No. 291, OECD Publishing, Paris, https://doi.org/10.1787/d62f618a-en.
"OpenAPI Specification," [Online]. Available: http://spec.openapis.org/oas/v3.0.3. [Accessed 24 03 2021].
"OpenAPI Specification v2.0," [Online]. Available: https://swagger.io/specification/v2/. [Accessed 24 03 2021].
OECD, Gennaio 2020. "The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector". Page 10–19. Retrieved January 2022, from: https://www.oecd.org/finance/TheImpact-Big-Data-AI-Insurance-Sector.pdf.
OXPER (2019). Building FinTech Ecosystems: Emerging Trends & Policy Implications. Insights from the 5th Annual Oxford Entrepreneurship Policy Roundtable (OXEPR). Available at: https://www.sbs.ox.ac.uk/sites/default/files/2020-09/OXEPR%202019%20%28Building%20FinTech%20Ecosystems%29.pdf.
Philippe Bracke, Anupam Datta, Carsten Jung and Shayak Sen, "Machine learning explainability in finance: an application to default risk analysis", Bank of England, Staff Working Paper No. 816, August 2019.
P. Samarati and L. Sweeney, "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression," 1998, [Online]. Available: http://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf.
Patricia Johnson, WhiteSource, "MiFID II Reforms and Their Impact on Technology and Security", Feb 2018, https://resources.whitesourcesoftware.com/blog

-whitesource/mifid-ii-reforms-and-their-impact-on-technology-and-security (accessed Nov 2020).

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., and Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. In Proceedings of Advances in Neural Information Processing Systems 31 (NeurIPS 2018).

Park, Hoyong, Eric Hsiao, and Andy Piper. "Continuous query language (CQL) debugger in complex event processing (CEP)." U.S. Patent No. 9,329,975. 3 May 2016.

P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, and K. Tzoumas, "Apache FlinkTM: Stream and Batch Processing in a Single Engine," p. 12.

P. Samarati and L. Sweeney, "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression," 1998, [Online]. Available: http://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf.

Pnevmatikakis, A., Kanavos, S., Matikas, G., Kostopoulou, K., Cesario, A., & Kyriazakos, S. (2021). Risk assessment for personalized health insurance based on real-world data. Risks. 9(3):46. https://doi.org/10.3390/risks9030046.

P. Treleaven, R. Gendal Brown and D. Yang, "Blockchain Technology in Finance," Computer, vol. 50, no. 9, pp. 14–17, 2017.

Pelleg, D., & Moore, A. (1999). Accelerating exact k-means algorithms with geometric reasoning. KDD '99: Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, 5(1), 277–281. ACM Digital Library. doi.org/10.1145/312129.312248.

Pentland, A. (2012). Society's Nervous System: Building Effective Government, Energy, and Public Health Systems. IEEE Computer, 45(1): 31–38.

Pinkas, B., and Lindell, Y. (2009). A proof of security of Yao's for two-party computation. J Cryptol 22, 161–188.

Pentland, A. (2014) Saving Big Data from itself. Scientific American, 65–68.

Perentis, C., Vescovi, M., Leonardi, C., Moiso, C., Musolesi, M., Pianesi, F., and Lepri, B. (2017). Anonymous or not? Understanding the factors affecting personal mobile data disclosure. ACM Trans. Internet Techn. 17(2): 13:1–13:19.

Pinkas, B., and Lindell, Y. (2009). A proof of security of Yao's for two-party computation. J Cryptol 22, 161–188.

Poursafaei, F., Hamad, G.B., & Zilic, Z. (2020). Detecting malicious Ethereum entities via application of machine learning classification. In Proceedings of the 2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS). IEEE.

P2957 – Standard for a Reference Architecture for Big Data Governance and Metadata Management. Standards.ieee.org. Retrieved January, 2022 from https://standards.ieee.org/project/2957.html.

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on information accompanying transfers of funds and certain crypto-assets (recast) COM/2021/422 final https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2021:422:FIN (Accessed: January 2022).

Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the mechanisms to be put in place by the Member States for the prevention of the use of the financial system for the purposes of money laundering or terrorist financing and repealing Directive (EU) 2015/849 COM/2021/423 final. EUR-Lex – 52021PC0423 – EN – EUR-Lex (europa.eu) (Accessed: January 2022).

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing COM/2021/420 final (Accessed: January 2022).

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing the Authority for Anti-Money Laundering and Countering the Financing of Terrorism and amending Regulations (EU) No 1093/2010, (EU) 1094/2010, (EU) 1095/2010 COM/2021/421 final https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0421 (Accessed: January 2022).

Pantelis Angelidis, Leslie Berman, Maria de la Luz Casas-Perez, Leo Anthony Celi, George E Dafoulas, Alon Dagan, Braiam Escobar, Diego M Lopez, Julieta Noguez, Juan Sebastian OsorioValencia, et al. (2016). The hackathon model to spur innovation around global mHealth. Journal of medical engineering & technology 40, 7–8, 392–399.

Pe-Than, E.P.P. et al. (2018): Designing Corporate Hackathons With a Purpose: The Future of Software Development. IEEE Softw. 36, 1.

"Pseudonymization," Imperva. https://www.imperva.com/learn/data-security/pseudonymization/.

PwC, "Redrawing the lines: FinTech's growing influence on Financial Services", Global FinTech Report 2017, available at; https://www.pwc.com/jg/en/publications/pwc-globalfintech-report-17.3.17-final.pdf.

P. M. N. Maló, "Hub-and-spoke Interoperability: an out of the skies approach for large-scale data interoperability," 2013, Accessed: Jan. 18, 2016. [Online]. Available: http://run.unl.pt/handle/10362/11397.

P. Andrews, I. Zaihrayeu, and J. Pane, "A classification of semantic annotation systems," Semantic Web, vol. 3, no. 3, pp. 223–248, 2012.

Proposal for a Regulation of the European parliament and of the council on Markets in Crypto-assets, and amending Directive (EU) 2019/1937, 2020, https://eu

r-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2020%3A0593%3AFIN.

P2957 – Standard for a Reference Architecture for Big Data Governance and Metadata Management. (n.d.). Standards.ieee.org https://standards.ieee.org/project/2957.html.

Policy and investment recommendations for trustworthy Artificial Intelligence. (2019, June 25). European Commission. Retrieved January 2022, from https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

Pidorycheva, I. Y. (2020). Innovation ecosystem in contemporary economic researches. Economy of Industry, (2 (90), 54–92.

Pascal Lafourcade, Marius Lombard-Platet, About blockchain interoperability, Information Processing Letters, Volume 161, 2020.

PReparing Industry to Privacy-by-design by supporting its Application in REsearch, Project Description, https://cordis.europa.eu/project/id/610613 (accessed in November 2020).

Presto – Distributed Query Engine for Big Data, https://prestodb.io/.

Priest, A. Na, M., Niedzwiadek, H. and Davidson, J. Sensor observation service. Technical Report OGC 06-009r6, October 2007.

Packer. [Online]. Available: https://www.packer.io/.

Q. Zhang, N. Koudas, D. Srivastava, and T. Yu, "Aggregate Query Answering on Anonymized Tables," in 2007 IEEE 23rd International Conference on Data Engineering, Apr. 2007, pp. 116–125.

Recommendations on outsourcing to cloud service providers, https://www.eba.europa.eu/sites/default/documents/files/documents/10180/2170121/5fa5cdde-3219-4e95-946d-0c0d05494362/Final%20draft%20Recommendations%20on%20Cloud%20Outsourcing%20%28EBA-Rec-2017-03%29.pdf?retry=1, [accessed March 2022].

Recital 19 EU GDPR, https://www.privacy-regulation.eu/en/recital-19-GDPR.htm, [accessed March 2022].

R. Chen, G. Acs, and C. Castelluccia, "Differentially private sequential data publication via variablelength n-grams," Proceedings of the 2012 ACM conference on Computer and communications security – CCS '12. 2012, doi: 10.1145/2382196.2382263.

R. Dewri, "Local Differential Perturbations: Location Privacy under Approximate Knowledge Attackers," IEEE Transactions on Mobile Computing, vol. 12, no. 12. pp. 2360–2372, 2013, doi: 10.1109/tmc.2012.208.

R. Cheng, Y. Zhang, E. Bertino, and S. Prabhakar, "Preserving User Location Privacy in Mobile Data Management Infrastructures," Privacy Enhancing Technologies. pp. 393–412, 2006, doi: 10.1007/11957454_23.

R. Castro Fernandez, M. Migliavacca, E. Kalyvianaki, and P. Pietzuch, "Making State Explicit for Imperative Big Data Processing," presented at the USENIX ATC'14: 2014 USENIX Annual Technical Conference, Philadelphia, USA, 2014.

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN.

Rancher, "Rancher overview," SUSE, [Online]. Available: https://rancher.com/docs/rancher/v2.5/en/overview/architecture/.

"RAML Version 1.0: RESTful API Modeling Language," [Online]. Available: https://github.com/raml-org/raml-spec/blob/master/versions/raml-10/raml-10.md/. [Accessed 24 03 2021].

R. M. M. M. M. A. Irakli Nadareishvili, Microservice Architecture: Aligning Principles, Practices, and Culture, O'Reilly Media, Inc., 2016.

Risks, an online journal, an international, scholarly, peer-reviewed, open access journal for research and studies on insurance and financial risk management. Risks is published monthly online by MDPI.

R. Campos-Rebelo, F. Moutinho, L. Paiva, and P. Maló, "Annotation Rules for XML Schemas with Grouped Semantic Annotations," in IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society, 2019, vol. 1, pp. 5469–5474.

Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. (2009). Bpr: Bayesian personalized ranking from implicit feedback. UAI '09: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, 452–461. ACM Digital Library. doi: 10.5555/1795114.1795167.

Rendle, S., Krichene, W., Zhang, L., & Anderson, J. (2020). Neural Collaborative Filtering vs. Matrix Factorization Revisited. RecSys '20: Fourteenth ACM Conference on Recommender Systems, 240–248. ACM DigitalLibrary. https://doi.org/10.1145/3383313.3412488.

Rendle, S., Zhang, L., & Koren, Y. (2019). On the difficulty of evaluating baselines: A study on recommender systems. arxiv Preprint, Information Retrieval. arXiv. https://arxiv.org/abs/1905.01395.

Rokach, L., & Maimon, O. (2005). Clustering methods. Data mining and knowledge discovery handbook. Springer US, 321–352.

Ruta, M., Noia, T. Di, Scioscia, F., Di Sciascio E. Semantic-enhanced EPCglobal Radio-Frequency IDentification. SWAP 2007.

RKE. [Online]. Available: https://rancher.com/docs/rke/latest/en/.

Rancher. [Online]. Available: https://rancher.com/.

R. S. M. S. W. M. &. S. R. Norvill, "Blockchain for the Simplification and Automation of KYC Result Sharing," in 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2019.

Rasmussen, C. E., Williams, C. K. I. (2006). Gaussian Processes for Machine Learning, The MIT Press.

R. Agrawal, C. Faloutsos, and A. N. Swami. E cient similarity search in sequence databases. In Proceedings of the International Conference on Foundations of Data Organization and Algorithms (FODO), pages 69–84. Springer-Verlag, 1993.

R. Cole, D. Shasha, and X. Zhao. Fast window correlations over uncooperative time series. In Proceedings of the International Conference on Knowledge Discovery and Data Mining (SIGKDD), pages 743–749. ACM, 2005.

Ruby. [Online]. Available: https://www.ruby-lang.org/en/.

Ribeiro, Marco & Singh, Sameer & Guestrin, Carlos. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. 1135–1144. 10.1145/2939672.2939778.

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN.

Rahimi, A., and Recht, B. (2008). Random features for large-scale kernel machines. In Proceedings of Advances in Neural Information Processing Systems 20.

"Random Forest Regressor", 2021 [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html#sklearn.ensemble.RandomForestRegressor.

R. J. Hyndman and G. Athanasopoulos. Forecasting: Principles and Practice. https://otexts.com/fpp2.

Richards, D., Rebeschini, P., and Rosasco L. (2020). Decentralised learning with random features and distributed gradient descent. In Proceedings of the 37th International Conference on Machine Learning (ICML), 119: 8105–8115.

Rissanen, E. (2013). Extensible Access Control Markup Language (XACML) Version 3.0. OASIS Standard. Available online at: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html.

S.-S. Ho and S. Ruan, "Differential privacy for location pattern mining," Proceedings of the 4th ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS – SPRINGL '11. 2011, doi: 10.1145/2071880.2071884.

Salinas, D., Flunkert, V., Gasthaus, J., Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting 36(3), 1181–1191.

Spark Streaming | Apache Spark., https://spark.apache.org/streaming/.

Stratosphere?» Next Generation Big Data Analytics Platform. http://stratosphere.eu/.

Samza., http://samza.apache.org/.

S. Schneider, H. Andrade, B. Gedik, A. Biem, and K. L. Wu, "Elastic scaling of data parallel operators in stream processing," in 2009 IEEE International Symposium on Parallel Distributed Processing, 2009, pp. 1–12.

Saft – How is the IoT world shaping up in 2021 and what trends will influence the future of IoT? https://www.saftbatteries.com/energizing-iot/how-iot-world-shaping-2021-and-what-trends-will-influence-future-iot-infographic.

Solera – Accelerating digital transformation for the future of automotive claims https://get.solera.com/SOLGLOBAL2020-04MTR1219Digital-Transformation-Webinar_05-Recording-Download.html.

S. Nexus, "Sonatype Nexus OSS," [Online]. Available: https://www.sonatype.com/nexus-repository-oss.

Slack, "Slack features," Microsoft, [Online]. Available: https://slack.com/intl/en-it/features.

Soldatos, John et al. "Big Data and Artificial Intelligence in Digital Finance: Increasing Personalization and Trust in Digital Finance Using Big Data and AI." Springer International Publishing AG, 2022.

Sonarqube, "Sonarqube architecture," SonarSource SA, 2020. [Online]. Available: https://docs.sonarqube.org/latest/architecture/architecture-integration/.

Spring Boot Framework, [Online]. Available: https://spring.io/projects/spring-boot. [Accessed 20 October 2020].

Spring Cloud Circuit Breaker, [Online]. Available: https://spring.io/projects/spring-cloudcircuitbreaker. [Accessed 21 October 2020].

Salesforce – What Is Digital Transformation? https://www.salesforce.com/products/platform/what-is-digital-transformation/.

Stack – Online guide to digital business transformation https://www.stackstudio.digital/onlineguide-to-digital-business-transformation/.

Shin, D. H., Kim, H., & Hwang, J. (2015). Standardization revisited: A critical literature review on standards and innovation. Computer Standards & Interfaces, 38, 152–157.

Self-Registration, 2020. [Online]. Available: https://microservices.io/patterns/self-registration.html. [Accessed 10 October 2020].

S. Assefa, D. Dervovic, M. Mahfouz, T. Balch, P. Reddy and M. Veloso, "Generating synthetic data in finance: opportunities, challenges and pitfalls," JPMorgan Chase & Co, 2020.

Springer (2021) "Digital Finance" Source: https://www.springer.com/journal/42521.

Source: https://www.fatf-gafi.org/publications/fatfrecommendations/documents/guidance-rba-virtual-assets2021.html.

Source: https://docs.google.com/spreadsheets/d/1_LghI9zgA5wGArjd1q6VeP7xWJzw1gx/edit?usp=sharing&ouid=115760979640495528344&rtpof=true&sd=true.

Su, Y. S., Zheng, Z. X., & Chen, J. (2018). A multi-platform collaboration innovation ecosystem: the case of China. Management Decision.

Supervision, B.: Basel committee on banking supervision. Principles for Sound Liquidity Risk Management and Supervision (September 2008) (2011).

Salehi, A., Aberer, K. «GSN, Quick and Simple Sensor Network Deployment», European conference on Wireless Sensor Networks (EWSN), Netherlands, 2007.

Scherp, A., Franz, T. Saatho, S. Staab. F–a Model of Events Based on the Foundational Ontology DOLCE+DnS Ultralight. In: International Conference on Knowledge Capturing (K-CAP), Redondo Beach, CA, USA., 2009.

Sharpe, W.F. (1994). The Sharpe Ratio. The Journal of Portfolio Management 21(1), 49–58.

Sheth, A. Henson, C., Sahoo. S. Semantic Sensor Web. IEEE Internet Computing 12 (4), 2008.

"Semantic Annotations for WSDL and XML Schema." https://www.w3.org/TR/sawsdl/ (accessed Jul. 22, 2020).

S. Peroni, "A simplified agile methodology for ontology development," in OWL: Experiences and Directions–Reasoner Evaluation, Springer, 2016, pp. 55–69.

S. A. McIlraith, T. C. Son, and H. Zeng, "Semantic web services," IEEE intelligent systems, vol. 16, no. 2, pp. 46–53, 2001.

S. Minukhin, V. Fedko and D. Sitnikov, "SQL-On-Hadoop Systems: Evaluting Performance of Polybase for Big Data Processing," 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018, pp. 591–594. [MIRAB12] Microsoft Industry Reference Architecture for Banking (MIRA-B), Microsoft Corporation Whitepaper, May 2012.

Shearer C., The CRISP-DM model: the new blueprint for data mining, J Data Warehousing (2000); 5:13—22. [Simitsis12] A. Simitsis, K. Wilkinson, M. Castellanos, U. Dayal, "Optimizing analytic data flows for multiple execution engines", in ACM SIGMOD, 2012, pp. 829–840.

Singh, A. (2019). Anomaly detection in the Ethereum network. A thesis for the degree of Master of Technology/Indian Institute of Technology Kanpur.

Sra, S., & Dhillon, I. (2005). Generalized Nonnegative Matrix Approximations with Bregman Divergences. Advances in Neural Information Processing Systems, 18(1), 283–290.

Stopar, L., Skraba, P., Grobelnik, M., & Mladenic, D. (2019). StreamStory: Exploring multivariate time series on multiple scales. IEEE Transactions on Visualization and Computer Graphics, 25(4), pp. 1788–1802.

Strumbelj, E., & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. Knowledge and Information Systems, 41(3), 647–665. Springer Link. doi.org/10.1007/s10115-013-0679-x.

Supporting Material – Archived Groups – Kantara Initiative, Kantarainitiative.org, 2020. [Online]. Available: https://kantarainitiative.org/confluence/display/archive/1+-+Supporting+Material. [Accessed 2 September 2020].

Sagi, O., and Rokach, O. L. (2018). Ensemble learning: A survey, WIREs Data Mining and Knowledge Discovery.

S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system", 2008. [Online] Available at: https://bitcoin.org/bitcoin.pdf [Accessed 19-October-2020].

Sen, J. (2013). Homomorphic encryption: Theory and applications. In J. Sen (ed.) Theory and practice of cryptography and network security protocols and technologies. INTECH Publishers.

Shamir, A. (1979). How to share a secret. Communications of the ACM, 22 (11): 612–613.

Staiano, J., Oliver, N., Lepri, B., de Oliveira, R., Caraviello, M., and Sebe, N. (2014). Money walks: A human-centric study on the economics of personal mobile data. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 583–594.

Staiano, J., Zyskind, G., Lepri, B., Oliver, N., and Pentland, A. (2019). The rise of decentralized personal data markets. In T. Hardjono, D. Shrier, and A. Pentland (eds.) Trust :: Data. A new framework for identity and data sharing. MIT Press.

Squirrel, SQL UI client, http://squirrel-sql.sourceforge.net/.

SPARQL Query Language for RDF Implementation Accessible here: http://www.w3.org/TR/rdf-sparql-query/.

The Financial Instrument Global Identifier https://www.openfigi.com/about/figi (Accessed: January 2022).

The Object Management Group OMG.org, an international, non-profit standards organization, founded in 1989.

The Financial Industry Business Ontology https://spec.edmcouncil.org/fibo/ (Accessed: January 2022).

The Financial Industry Business Ontology https://spec.edmcouncil.org/fibo/ (Accessed: January 2022).

The European project for Standardized Transparent Representations in order to Extend Legal Accessibility (Estrella, IST-2004-027655) http://www.estrellaproject.org.

The European project for Standardized Transparent Representations in order to Extend Legal Accessibility (Estrella, IST-2004-027655) http://www.estrellaproject.org.

"The EU regulation on electronic identification and certification services", 2016, https://www.dpc.bg/p/doc/dpc-dpco-the-eu-regulation-on-electronic-identification-and-certification-services-paving-the-way-forward-towards-more-secure-internet-transactions-07-04-2016-637.pdf, accessed Nov 2020.

"Tokenization Market by Component, Application Area (Payment Security, Application Area, and Compliance Management), Tokenization Technique (API-based and Gateway-based), Deployment Mode, Organization Size, Vertical, and Region – Global Forecast to 2023", 2018. MarketsandMarkets Research Report, [Online]. Available at: https://www.marketsandmarkets.com/MarketReports/tokenization-market76652221.html?gclid=Cj0KCQjw2or8BRCNARIsAC_ppyY4t6bLXgBQG1pregsjn3hKtxG8goxXR6RgjRXxlP1H7MY2IXUSqGsaAjNAEALw_wcB [Accessed: 20-October-2020].

Text Retrieval conference https://trec.nist.gov/ (Accessed: January 2022).

The Power MBA – The Lean Startup Methodology Cycle: 4 Steps to Risk-Free Success https://www.thepowermba.com/en/blog/lean-startup-methodology.

The Viima Blog – Innovation Management https://www.viima.com/blog/innovationmanagement?hs_amp=true.

The Viima Blog – Innovation Management – Processes https://www.viima.com/blog/innovation-management-processes?hs_amp=true.

The EC blockchain strategy in a nutshell, https://ec.europa.eu/digital-single-market/en/blockchaintechnologies (Accessed: January 2022).

The Enterprisers project – What is digital transformation https://enterprisersproject.com/sites/default/files/what_is_digital_transformation_2020.pdf.

Testing two-factor authentication, Michelangelo van Dam, 05/08/2017. https://www.in2it.be/2017/08/testing-two-factor-authentication/.

The KTH Innovation Readiness Level https://kthinnovationreadinesslevel.com/take-a-tour/.

Team Readiness Capability Assessment Model https://teamreadiness.com/?page_id=122.

T. Yuanyuan, T. Zou, F. Özcan, R. Gonscalves, H. Pirahesh, "Joins for hybrid warehouses: exploiting massive parallelism in hadoop and enterprise data warehouses", in EDBT/ICDT Conf., pp. 373–384 (2015).

T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," Scientific american, vol. 284, no. 5, pp. 34–43, 2001.

Tapanainen, T. (2020). Toward Fintech Adoption Framework for Developing Countries-A Literature Review based on the Stakeholder Perspective. Journal of Information Technology Applications and Management, 27(5), 1–22.

Tsiatsis, V., Gluhak, A., Bauge, T., Montagut, F., Bernat, J., Bauer, M., Villalonga, C., Barnaghi, P.M., Krco, S. The SENSEI Real World Internet Architecture. Future Internet Assembly, IOS Press, 2010.

Team, C. (2020). Indirect exposure: Why you need to look beyond direct counterparties to understand cryptocurrency address risk. Nov. 2020, https://blog.chainalysis.com/reports/cryptocurrency-risk-blockchain-analysis-indirect-exposure.

T. Cameron McKenzie, "open API (public API)," [Online]. Available: https://searchapparchitecture.techtarget.com/definition/open-API-public-API. [Accessed 24 03 2021].

T. Slimani, "Semantic annotation: The mainstay of semantic web," arXiv preprint arXiv:1312.4794, 2013.

T. M. Siebel, Digital transformation: survive and thrive in an era of mass extinction. RosettaBooks, 2019.

Thompson, Gregory, et al. "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization." Monthly Weather Review 136.12 (2008): 5095–5115.

«Terraform,» [Online]. Available: https://www.terraform.io/.

«Terraform homepage,» HashiCorp, [Online]. Available: https://www.terraform.io/docs/language/files/index.html.

T. Özsu, P. Valduriez, Principles of Distributed Database Systems, 4th ed. Springer, 700 pages (2020).

Takemori, S., and Sato, M. (2020). Approximation methods for kernelized bandits, arXiv 2010.12167.

Tamer Özsu, Patrick Valduriez. Principles of Distributed Database Systems, 4th Edition, Springer, 2020.

UN AHEG (2020), First version of a draft text of a recommendation on the ethics of Artificial Intelligence, SHS/BIO/AHEG-AI/2020/4, limited distribution.

Understanding JSON Schema, [Online]. Available: http://json-schema.org/understanding-json-schema/. [Accessed 24 03 2021].

Unified Medical Language System Accessible here: http://www.nlm.nih.gov/research/umls/index.html.

UN AHEG (2020), First version of a draft text of a recommendation on the ethics of Artificial Intelligence, SHS/BIO/AHEG-AI/2020/4, limited distribution.

Upbit hacking incident announcement, Nov. 2019, https://upbit.com/service_center/notice?id=1085.

Verbatim text from https://resources.whitesourcesoftware.com/legal/mifid-ii-reforms-and-their-impact-on-technology-and-security (accessed in December 2020).

V. Ayala-Rivera, P. McDonagh, T. Cerqueus and L. Murphy, "Synthetic Data Generation using Benerator Tool," University College Dublin, 2013.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is All you Need. Advances in Neural Information Processing Systems, 30, pp. 5998–6008.

V. Buterin, "A next-generation smart contract and decentralized application platform, Ethereum Project White Paper", [Online] Available at: https://github.com/ethereum/wiki/wiki/White-Paper [Accessed: 20-October-2020].

Vescovi, M., Lepri, B., Perentis, C., Moiso, C., and Leonardi, C. (2014). My data store: Toward user awareness and control on personal data. In Proceedings of UbiComp Adjunct: 179–182.

Verbatim text from https://resources.whitesourcesoftware.com/legal/mifid-ii-reforms-and-their-impact-on-technology-and-security (accessed in December 2020).

Vagrant. [Online]. Available: https://www.vagrantup.com/.

VMware. [Online]. Available: https://www.vmware.com/.

v. ESXi. [Online]. Available: https://www.vmware.com/products/esxi-and-esx.html.

«VMware Workstation,» [Online]. Available: https://www.vmware.com/products/workstationpro.html.

VMware. [Online]. Available: https://www.vmware.com/.

V. Buterin, Chain interoperability, https://www.r3.com/wp-content/uploads/2017/06/chain_interoperability_r3.pdf (Accessed: January 2022).

V. Gadepally, P. Chen, J. Duggan, A. J. Elmore, B. Haynes, J. Kepner, S. Madden, T. Mattson, M. Stonebraker, "The BigDawg polystore system and architecture", in IEEE High Performance Extreme Computing Conference (HPEC), 2016, pp. 1–6.

V. Gulisano, R. Jiménez-Peris, M. Patiño-Martínez, C. Soriente and P. Valduriez, "StreamCloud: An Elastic and Scalable Data Streaming System," in IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 12, pp. 2351–2365, Dec. 2012.

WSO2 CEP., https://wso2.com/products/complex-event-processor/.

WITDOM Project Documentation https://bit.ly/3cEA75e.

Web Services Description Language (WSDL) 1.1, [Online]. Available: https://www.w3.org/TR/2001/NOTE-wsdl-20010315. [Accessed 24 03 2021].

Web Application Description Language, [Online]. Available: https://www.w3.org/Submission/wadl/. [Accessed 24 03 2021].

WBCINNO Project – Methodology for Innovation Management, University of Kragujevac, Prof. Dr. Slobodan Arsenijević http://www.wbc-inno.kg.ac.rs/pub/download/13953889933298_wbc_inno_academic_methodology_for_innovation_management_eng.pdf.

WBC Inno – Methodology for Innovation Management http://www.wbc-inno.kg.ac.rs/pub/download/13953889933298_wbc_inno_academic_methodology_for_innovation_management_eng.pdf.

Wikipedia, "Bastion Host," [Online]. Available: https://en.wikipedia.org/wiki/Bastion_host.

World Wide Web Consortium – Linked Open Data, Accessible here: http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData.

World Wide Web Consortium – Resource Description Framework Schema Accessible here: http://www.w3.org/TR/rdf-schema.

World Wide Web Consortium – Turtle Serialisation Specification Accessible here: http://www.w3.org/TeamSubmission/turtle/.

World Wide Web Consortium – N-Triples format specification Accessible here: http://www.w3.org/TR/rdf-testcases/#ntriples.

World Wide Web Consortium – Ontology Web language Accessible here: http://www.w3.org/TR/owl-ref.

World Wide Web Consortium – Resource Description Framework in Attributes Accessible here: http://www.w3.org/TR/xhtml-rdfa-primer/.

Wang, X., He, X., Wang, M., Feng, F., & Chua, T.-S. (2019). Neural Graph Collaborative Filtering. SIGIR'19: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 165–174. ACM digital Library. https://doi.org/10.1145/3331184.3331267.

W. Entriken, D. Shirley, E. Evans, and N. Sachs, "ERC-721 non-fungible token standard," 2018. [Online]. Available: https://eips.ethereum.org/EIPS/eip-721. [Accessed 1-October-2020].

Want, R., Pering, T., Danneels, G., Kumar, M., Sundar, M., and Light, J. (2002). The personal server: Changing the way we think about ubiquitous computing. In Proceedings of 4th International Conference on Ubiquitous Computing, 194–209.

What is a lambda architecture? https://databricks.com/glossary/lambda-architecture.

World Economic Forum Companion to the Model AI Governance Framework—Implementation and Self-Assessment for Organizations (Infocomm Media Development Authority of Singapore, 2020) Retrieved January 2022 from,

https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGIsago.pdf.

Will Kenton and Julius Mansa, "MiFID II – Laws&Regulations", Jul 2020, https://www.investopedia.com/terms/m/mifid-ii.asp (Accessed Nov 2020).

Wedge R., Kanter J. M., Veeramachaneni K., Rubio S. M., and Perez S. I., (2018). Solving the false positives problem in fraud prediction using automated feature engineering. In Proceedings of the.

W. Radomski, A. Cooke, P. Castonguay, J. Therien, E. Binet, and R. Sandford, "ERC-1155 multi token standard," 2015. [Online]. Available: https://eips.ethereum.org/EIPS/eip-1155 [Accessed 10-October-2020].

W. Radomski, A. Cooke, P. Castonguay, J. Therien, E. Binet, and R. Sandford, "ERC-1155 multi token standard," 2015. [Online]. Available: https://eips.ethereum.org/EIPS/eip-1155 [Accessed 10-July-2021].

Y. Liao, M. Lezoche, H. Panetto, and N. Boudjlida, "Semantic annotation model definition for systems interoperability," in OTM Confederated International Conferences "On the Move to Meaningful Internet Systems", 2011, pp. 61–70.

Y. Song, X. Lu, S. Nobari, S. Bressan, and P. Karras, "On the privacy and utility of anonymized social networks," International Journal of Adaptive, Resilient and Autonomic Systems (IJARAS), vol. 4, no. 2, pp. 1–34, 2013.

Y.-A. de Montjoye, Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the Crowd: The privacy bounds of human mobility," Scientific Reports, vol. 3, no. 1. 2013, doi: 10.1038/srep01376.

Yagoubi, D.-E., Akbarinia, R., Kolev, B., Levchenko, O., Masseglia, F., Valduriez, P., Shasha, D., 2018. ParCorr: Efficient Parallel Methods to Identify Similar Time Series Pairs across Sliding Windows. Data Mining and Knowledge Discovery, vol. 32(5), pp. 1481–1507. Springer.

Yang, L., Bagdasaryan, E., Gruenstein, J., Hsieh, C.-K., & Estrin, D. (2018). Openrec: A modular framework for extensible and adaptable recommendation algorithms. WSDM '18: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 664–672. ACM Digital Library. https://doi.org/10.1145/3159652.3159681.

Z. Minpeng, R. Tore, "Querying combined cloud-based and relational databases", in Int. Conf. on Cloud and Service Computing (CSC), 2011, pp. 330–335.

Zyskind, G., Nathan, O., and Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. In Proceedings of IEEE Symposium on Security and Privacy Workshops: 18.

Zeng, S., Tay, Y., Yao, L., Wu, B., & Sun, A. (2019). Deeprec: An open-source toolkit for deep learning based recommendation. Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, 6581–6583. dblp computer science bibliography. https://doi.org/10.24963/ijcai.2019/963.

Zyskind, G., Nathan, O., and Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. In Proceedings of IEEE Symposium on Security and Privacy Workshops: 180–184.

"3rd-Party Registration," 2020. [Online]. Available: https://microservices.io/patterns/3rd-party-registration.html. [Accessed 15 October 2020].