PART I:

CONCEPTS AND DESIGN THINKING INNOVATION ADDRESSING THE GLOBAL FINANCIAL NEEDS

THE INFINITECH WAY FOUNDATIONS

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: 978-1-63828-228-0
E-ISBN: 978-1-63828-229-7
DOI: 10.1561/9781638282297

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 1. Concepts and Design Thinking Innovation Addressing the Global Financial Needs: The INFINITECH Way Foundations*. 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 viii

Who Should Read This Book? x

What is Addressed in the Book Series? xi

What is Covered in this INFINITECH Part I Book? xiii

Acknowledgements xv

Contributing Authors xvi

Abstract xxv

Chapter 1 INFINITECH and The Global Financial Sector 1

1.1 INFINITECH and The Global Financial Sector 1

Chapter 2 INFINITECH Way Foundations 6

2.1 INFINITECH Way Foundations 6
2.2 The INFINITECH Way Innovation 8
2.3 Innovation Radar 10
2.3.1 Customer 11
2.3.2 Market 11
2.3.3 Compliance 14
2.3.4 Technology 15
2.3.5 Resources 18
2.4 Certification 20
2.5 Infinitech Terminology 21

Chapter 3 Reference Architectures Analysis 22
3.1 Reference Architecture Analysis 22
3.2 Methodology 23
3.3 General AI/BigData Challenges 24
3.3.1 Big Data 24
3.3.2 Data Pipelines 27
3.4 Specific Challenges for the Finance Sector 28
3.4.1 Siloed Data and Business Operations 28
3.4.2 Real Time Performance Requirements 28
3.4.3 Mobility 29
3.4.4 Omni-Channel Banking – Multiple Channels Management 29
3.4.5 Automation 29
3.4.6 Transparency 30
3.5 Reference Architecture for Big Data/IoT/AI in Finance/Insurance . 30
3.6 INFINITECH Reference Architecture 31
Chapter 4 INFINITECH Data Pack 38
4.1 INFINITECH Data Pack 38
4.1.1 FIBO, Financial Industry Business Ontology 39
4.1.2 FIGI Financial Instrument Global Identifier 39
4.1.3 LKIF, the Legal Knowledge Interchange Format 40
4.1.4 INFINITECH Core 40
4.1.5 Example-TAHO (Traffic Analysis Hub Ontology) 43
4.1.6 Terminology Used in INFINITECH 46
Chapter 5 INFINITECH Technologies, Data and Processes 49
5.1 INFINITECH Technologies, Data, and Processes 49
5.2 Technologies 50
5.2.1 Data Ingestion 50
5.2.2 Data Protection Orchestrator (DPO) 50
5.2.3 Digital User Onboarding System (DUOS) 51
5.2.4 EASIER-AI 51
5.2.5 Driver Profile Classifier 51
5.2.6 Distributed Near-real-time HPC Processing and Exchange of IoT Streaming Data 51
5.2.7 Botakis Chatbot Development Network 52
5.2.8 Crowdpolicy Open (Innovation) Banking Solution 52
5.2.9 AI-Engine-for-Psychometric-Profiling and Personalization 52
5.2.10 Open Source AI/ML Frameworks 52
5.2.11 Data Layer – REST API 53

5.2.12 Terrier Information Retrieval Platform 53
5.2.13 Anonymization Tool 53
5.2.14 Polyglot Database Management System 53
5.2.15 HTAP Database for the Financial and Insurance Sector 54
5.2.16 Natural Language Processing for Real-time, High-accuracy Credit Risk Assessment 55
5.2.17 Machine Learning Algorithms for Health Related Data 55
5.2.18 Wenalyze Big Data Analytics Platform 55
5.2.19 Octopush Geospatial Enabling Framework 55
5.2.20 AgroApps Weather Intelligence Engine 56
5.2.21 Sentiment Analysis Tool 56
5.2.22 Partitioned and Distributed Transaction Graphs 57
5.2.23 ALIDA: A Micro-service Based Platform for Composition, Deployment and Execution of BDA Applications 57
5.2.24 Text Analysis Tool 57
5.2.25 Blockchain Tokenization 57
5.2.26 Healthentia LifeSciences BigData Platform 57
5.2.27 Event-registry 58
5.2.28 Stream Story 58
5.2.29 Qminer 58
5.2.30 SSC – Super Stream Collider, a Multiformat Data Management and Query System 58
5.2.31 The Global Engine with Neural Network Intelligence (GENNI) . . 59
5.2.32 Data Check-in Mechanism 59
5.2.33 IoT-Catalogue 59
5.2.34 Analytics Library 60
5.2.35 Catalogue of Objects 60
5.2.36 Portfolio Optimization Tool 60
5.2.37 Sentiment Analysis for Financial News 60
5.2.38 AI Model for VaR Prediction 61
5.2.39 INFINITECH Components Specifications 61
5.2.40 Data Component Description 61
5.3 Data Components 62
5.3.1 Relational Database 62
5.3.2 Polyglot Query Processing 62
5.3.3 Incremental Analytics 63
5.3.4 OneHotEncoder 63
5.3.5 Stream Processor 64
5.3.6 Online Aggregates 64

5.3.7 Data Collection 64
5.3.8 Anonymization Tool 65
5.3.9 DUOS (Digital User On-boarding Services) 65
5.3.10 DPO (Data Protection Orchestrator) 65
5.3.11 Blockchain Reader 66
5.3.12 Blockchain Writer 66
5.3.13 Smart Contract Executor 66
5.3.14 Blockchain Data Visualizer 66
5.3.15 Blockchain Authenticator 66
5.3.16 Blockchain Encryptor 67
5.3.17 Blockchain Decryptor 67
5.3.18 Blockchain Transaction Dataset Preparation Component 67
5.3.19 Scalable Transaction Graph Analysis Component 67
5.3.20 Semantic Streams Analyzer 67
5.3.21 Semantic Reasoner 68
5.3.22 Ontology Mapping 68
5.3.23 Semantic Annotator-Preprocessing 69
5.3.24 Smart Fleet (IoT Context Management and Historical Data Component) 69
5.3.25 Fraud Detection Service Training 69
5.3.26 Fraud Detection Service Execution 69
5.3.27 Pay As You Drive Service Training 70
5.3.28 Pay As You Drive Service Execution 70
5.3.29 Investment Recommendation Engine Training 70
5.3.30 Investment Recommendation Engine Execution 70
5.3.31 Recommender 70
5.3.32 Cash Flow Prediction 70
5.3.33 Budget Prediction 71
5.3.34 KPI Engine 71
5.3.35 Transaction (Txn) Monitoring 71
5.3.36 Transaction (Txn) Categorization 71
5.3.37 Invoice Processing 71
5.3.38 KMeans 72
5.3.39 Random Forest (Model) 72
5.3.40 Random Forest (Predict) 72
5.3.41 Client Contextual Information 72
5.3.42 Financial Fraud/Crime Risk Score 72
5.3.43 Anomaly Analysis 72

5.3.44 Pattern Analysis 73
5.3.45 Stream Story 73
5.3.46 Open API Gateway 73
5.3.47 User Interface for Blockchain Transaction Reports and Visualization Component 73
5.3.48 Visualization Preparation 73
5.3.49 Real Time Visualization 74
5.3.50 INFINISTORE 74
5.3.51 UI Risk Assessment Based on VaR 74
5.3.52 Pseudo-anonymization Tool 74
5.3.53 Health Insurance Risk Assessment Service 74
5.3.54 Health Insurance Fraud Detection Service 74
5.3.55 Well-being Outlook Classifiers 74
5.3.56 Synthetic RWD for Well-being Analytics 75
5.3.57 Open Banking Agreggator Solution 75
5.3.58 Big Data Analytics Platform 75

5.4 Processes 75
5.4.1 Semantics Streams Analytics Engine (SeSA-ME) 75
5.4.2 CI/CD 76
5.4.3 KYC/KYB 77

5.5 Self-assessment 78

Chapter 6 INFINITECH Way Foundations Impact on Fintech and Insurance 81

6.1 INFINITECH Way Foundations Impact on Fintech and Insurance 81
6.2 The Process of Adopting INFINITECH Way Foundations 93

Chapter 7 Conclusions 95

7.1 Conclusions 95

Chapter 8 References 98

8.1 Infinitech Public Deliverables 98
8.2 References 102

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.

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 I Book?

“Concepts and Design Thinking Innovation addressing the Global Financial Needs”

The INFINITECH Way Foundation is explained in simple words, introducing the design principles and basic metrics for enhancing the state of the art. In this book we illustrate how INFINITECH Way can offer many advantages in the on-boarding process offering speed and agility when optimizing the design and implementation of financial services reference architecture and/or without the problems of dealing with complex vendor-lock and/or proprietary infrastructure.

It is also addressed how INFINITECH Way is very practical in multiple areas and how the INFINITECH Way is applied. The core elements of the INFINITECH Way is the Reference Architecture (RA) and the different methods to assess the success of the implementation. Further, it is shown how the INFINITECH Way is applied to validate a Reference Implementation addressing Big Data, IoT and AI applications for the financial and insurance sectors. The reference implementation will serve as a blueprint for the rapid and cost-effective solutions development and deployment. The INFINITECH-RA classify and describe a set of common functional building blocks that will support advances of Big Data, AI and IoT applications.

In this book, a concise and detailed overview of the INFINITECH Reference Architecture is provided. The INFINITECH Data Pack is also explained containing the set of files, schemas, and metadata model diagrams (Graphs) that represent how the INFINITECH way and the financial data could be organized and structured. Furthermore, the core vocabularies of the INFINITECH Data Pack including FIBO, FIGI, LKIF, and INFINITECH Core are cited and explained. This book also includes a practical example of how the INFINITECH Core is used and describes possible extensions in terms of a global technology named traffic analysis Hub and its work towards formalizing the vocabularies as the TAHO (Traffic Analysis Hub Ontology). Finally, a quick run-down of INFINITECH technologies, data, and processes applied in the project are listed. In the section on self-assessment, we describe its integration with RA. INFINITECH WAY impact on Fintech and Insurance is discussed and its application to pilots is illustrated.

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

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

Abstract

In this 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 envisions a large number of 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.

The INFINITECH Way Foundations is explained in simple words, introducing the design principles and basic metrics for enhancing the state of the art. In this book we illustrate how INFINITECH Way Foundations can offer many advantages in the on-boarding process offering speed and agility when optimizing the design and implementation of financial services reference architecture and/or without the problems of dealing with complex vendor-lock and/or proprietary infrastructure. It is also illustrated how INFINITECH Way Foundations is very practical in multiple areas and how the INFINITECH Way Foundations is applied. The core elements of the INFINITECH Way Foundation are the Reference Architecture (RA) and the different methods to assess the success of the implementation. Further, it is shown how the INFINITECH Way Foundations are applied to validate a Reference Implementation addressing BigData, IoT and AI applications for the financial and insurance sectors. The reference implementation will serve as a blueprint for the rapid and cost-effective solutions development and deployment. The INFINITECH-RA classifies and describes a set of common functional building blocks that will support advances of BigData, AI and IoT applications.

In this book, a concise and detailed overview of the INFINITECH Reference Architecture is provided. The INFINITECH Data Pack is also explained containing the set of files, schemas, and metadata model diagrams (Graphs) that represent how the INFINITECH Way Foundations and the financial data could be organized and structured. Furthermore, the core vocabularies of the INFINITECH Data Pack including FIBO, FIGI, LKIF, and INFINITECH Core are cited and explained. This book also includes a practical example of how the INFINITECH Core is used and describes possible extensions in terms of a global technology named traffic analysis Hub ad its work towards formalizing the vocabularies as the TAHO (Traffic Analysis Hub Ontology). Finally, a quick run-down of INFINITECH technologies, data, and processes applied in the project are listed. In the section on self-assessment, we describe its integration with RA. INFINITECH Way Foundations impact on Fintech and Insurance is discussed and its application to pilots is illustrated.

DOI: 10.1561/9781638282297.ch1

Chapter 1

INFINITECH and The Global Financial Sector

1.1 INFINITECH and The Global Financial Sector

The finance sector is among the most data-savvy and data-intensive of the global economy. The on-going digital transformation of financial organizations, along with their interconnection as part of a global digital finance ecosystem is producing petabytes of structured and unstructured data. The latter represent a significant opportunity for banks, financial institutions, and financial technology firms (FinTechs): Leveraging these data financial organizations can significantly improve both their business processes and the quality of their decisions. As a prominent example, modern banks can exploit customer data to anticipate the behaviors of their customers, and to deliver personalized banking solutions to them. Likewise, data can enable new forms of intelligent algorithmic trading and personalized asset management [Soldatos *et al.*, 2022].

To harness the benefits of big data, financial organizations need effective ways for managing and analyzing large volumes of structured, unstructured, and semi-structured data at scale. Furthermore, they need to manage both streaming data and data at rest, while at the same time providing the means for the scalable processing by a variety of analytics algorithms. The latter processing may also require support for real-time analytics over the heterogeneous types of data (i.e. structured/unstructured, streaming/data-at-rest). Management of heterogeneity is therefore one of the most important concerns for big data management in the financial sector. Currently, financial organizations spend significant effort and IT resources in unifying and processing different types of data, which typical reside in different repositories such as operational databases, analytical data bases, data warehouses, data lakes, as well as emerging distributed ledger infrastructures (i.e. blockchains). Moreover, several applications require also semantic interoperability across datasets that are "siloed" across different systems, while the use of big data in real-life banking applications also requires the utilization of pre-processing functions (e.g., anonymization), which boost the ever-important regulatory compliance of digital finance applications. Despite the evolution of big data technologies, there is still a need for novel solutions that can successfully confront the above-listed challenges to enable the development, deployment, and operation of big data applications for digital finance at scale [Soldatos *et al.*, 2022].

Big data management and analytics solutions enable Artificial Intelligence (AI) solutions in finance. AI refers to the capability of machines to imitate intelligent human behavior and act like humans. In many cases AI systems think like humans and reason over complex contexts in order to evaluate and to take optimal decisions. As such, AI systems support two main processes: (i) A learning process that allows them to produce rules about how to process and use the information they receive, and (ii) Reasoning that drives their decisions and actions. Several AI systems are already deployed or planned to be integrated in applications of the financial services industry. They are usually based on one or a combination of technologies such as video processing and visual scene analysis, speech recognition, Natural Language Processing (NLP), automated translation, machine learning, deep learning, and cognitive search. Typical AI applications in digital finance include robo-advisors, AI-based personalized asset management systems, statistical credit underwriting and risk assessment applications, automated and intelligent KYC (Know Your Customer) applications, fraud detection and anti-money laundering (AML), personalized finance applications for citizens and businesses, as well as a variety of front office applications such as chatbots. Moreover, there are many interesting AI applications in the insurance sector such as automated insurance claims management and usage-based insurance i.e., statistical calculation of risk premiums based on data about the customers' behaviors (e.g., lifestyle or driving behavior data). Most of these use cases leverage Machine Learning (ML) and Deep Learning (DL) techniques. However, the above list of AI use cases in finance is non-exhaustive. As more data becomes available, the use of AI to improve automation, personalization and reduce costs will become more attractive. It is expected that FinTech enterprises will produce novel AI-based ideas in the years to come. Nevertheless, in several cases AI deployments have to overcome barriers and limitations of existing big data management technologies. In other cases, integration with other emerging technologies (e.g., RPA – Robotics Process Automation and blockchain technologies) are required. In this context, the presentation of tangible AI deployments in financial institutions is interesting in terms of the technologies employed, as well as in terms of the integration of AI in the digital transformation strategies of the financial organizations [Soldatos *et al.*, 2022].

The regulatory compliance of big data and AI solutions is also a major concern of financial institutions. Recent regulatory developments in the finance sector, such as the 2nd Payment Services Directive (PSD2) in Europe, open significant innovation opportunities for the sector, through facilitating the flow of data across financial organizations. At the same time, these regulations make compliance processes more challenging, while introducing security challenges. Other regulatory developments (e.g., MiFIDII and the 4th AML Directive) affect the way certain use cases (e.g., fraud detection) are developed. Moreover, all data-intensive use cases should comply with data protection regulations such as the General Data Privacy Regulation (GDPR) of the European Union (EU). At the same time, there are new regulations (e.g., eIDAS) which facilitate certain tasks in large scale big data and AI use cases, such as the process of digital on-boarding and verification of customers. Overall, regulatory compliance has a two-way interaction with big data and AI applications. On the one hand it affects the design and deployment of big data and AI applications, while on the other big data and AI can be used to boost regulatory compliance. Indeed, AI provides many opportunities for improving regulatory compliance in the direction of greater accuracy and cost-effectiveness. For instance, the machine learning segment of AI enables collection and processing of very large amounts of data relevant to a financial or banking workflow, including structured data from conventional banking systems and alternative data such as news and social networks. These big data can be analyzed to automate compliance against regulatory rules. This is the reason why many RegTech (Regulatory Compliance) enterprises are AI-based. They leverage AI and big data analytics to audit compliance in real time (e.g., by processing streaming data in real time), while at the same boosting the accuracy and richness of regulatory reporting.

Overall, the development, deployment and operation of novel big data and AI solutions for modern digital financial organization requires a holistic approach that addresses all the above listed issues. Since October 2019, the European project INFINITECH (co-funded by the H2020 program of the European Commission) is taking such a holistic and integrated approach to designing, deploying and demonstrating big data and AI solutions in Digital Finance. The project brings together a consortium of over forty organizations, including financial organizations, FinTechs, large vendors of AI and big data solutions, innovative high-tech Small

Medium Enterprises (SMEs), as well as established research organizations with a proven track record of novel research outcomes in AI, big data and blockchain technologies and their use in the finance sector. The present book is aimed at presenting INFINITECH's approach to big data and AI driven innovations in Digital Finance, through a collection of scientific and technological development contributions that addresses most of the earlier identified challenges based on innovative solutions. Specifically, the book presents a set of novel big data, AI and blockchain technologies for the finance sector, along with their integration in novel solutions. Furthermore, it puts emphasis on regulatory compliance issues, including technological solutions that boost compliance to some of the most important regulations of the sector.

INFINITECH is a joint effort of Europe's leaders in ICT and Finance/Insurance sectors towards providing the technological capabilities, the experimentation facilities (testbeds & sandboxes) and the business models needed to enable European financial organizations, insurance enterprises and FinTech/InsuranceTech innovators to fully leverage the benefits of BigData, IoT and AI technologies. The latter benefits include a shift towards autonomous (i.e. automated and intelligent) processes, that are dynamically adaptable and personalized to end-users' needs, while being compliant to the sector's complex regulatory environment. It will provide: [the proposal]

- Novel Big Data/IoT technologies for seamless management and querying of all types of data (e.g., OLAP/OLTP, structured/ unstructured/semi-structured, data streaming & data at rest), interoperable data analytics, blockchain-based data sharing, real-time analytics, as well as libraries of advanced AI algorithms.
- Regulatory tools incorporating various data governance capabilities (e.g., anonymization, eIDAS integration) and facilitating compliance to various regulations (e.g., PSD2 2nd Payment Services Directive), 4AMLD (4th Anti-Money Laundering), MIFiD II).
- Novel and configurable testbeds & sandboxes, each one offering Open APIs and other resources for validating autonomous and personalized solutions, including a unique collection of data assets for finance/insurance.

The project's results will be validated in the scope of 15 high impact pilots providing complete coverage of the sectors, including Know Your Customer (KYC), customer analytics, personalized portfolio management, credit risk assessment, preventive financial crime analysis, fraud anticipation, usage-based insurance, agro-insurance and more. INFINITECH will 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. Based on their strong footprint in the European digital finance ecosystem, the partners will engage stakeholders from all EU-28 countries, making INFINITECH synonymous to disruptive BigData/AI innovation in the target sectors.

DOI: 10.1561/9781638282297.ch2

Chapter 2

INFINITECH Way Foundations

2.1 INFINITECH Way Foundations

The INFINITECH Project has introduced the “INFINITECH Way Foundations” which offers the best practices in FinTech developments for the insurance and financial sectors in Europe. [Innovation Readiness Assessment]

INFINITECH has introduced the “INFINITECH Way Foundations” which offers many advantages in the on-boarding process offering speed and agility when optimizing the reference architecture and without the problems of dealing with complex, proprietary infrastructure. This on-boarding process outlines adopting containers and orchestration and transitioning implemented solutions to a containerized approach. INFINITECH Readiness Level (IRL) focuses on evaluating the paths taken in the adoption and deployment of those paths and technologies for an easy adoption and efficient deployment. [Innovation Readiness Assessment]

The INFINITECH project has successfully designed and used a holistic framework called INFINITECH Business Approach or “INFINITECH Way Foundations” to navigate through the design, development, and deployment of technology solutions in the financial services and insurance sectors. [INFINITECH Project Review Report]

The resulting “INFINITECH Way Foundations” provides a strong technical framework for developing data driven financial services related applications. The business case for using the technical framework, for example, through a business model specification, has not (yet) been validated. By relying on the "INFINITECH Way Foundations" as a development framework SMEs, notably FinTech and InsuranceTech companies should be able to advance more effectively and avoid reinventing the wheel over and over. [INFINITECH Project Review Report]

In the other hand INFINITECH is using an innovative on-boarding called the "INFINITECH Way Foundations" which also act as a methodology outlining containerized solutions and orchestrated Development and Operations (DevOps) operations and transitioning methods for implemented software assets, components, and solutions. For example, the "INFINITECH Way Foundations" is an innovative product that provides a strong technical framework for developing data driven financial services related applications, but a strong business model has not been clearly specified. [Innovation Readiness Assessment]

INFINITECH Way Foundations applies to:

- INFINITECH Reference Architecture (RA)
- INFINITECH Data pack
    - FIBO, Financial Industry Business Ontology
    - FIGI Financial Instrument Global Identifier
    - LKIF, the Legal Knowledge Interchange Format
    - INFINITECH Core
    - Example-TAHO (Traffic Analysis Hub Ontology)
- INFINITECH data, technologies, and processes
    - Interoperability Data Pack
    - CI/CD
    - KYC/KYB
- INFINITECH Way Self-assessment
- INFINITECH Way impact on Fintech and Insurance

Overall, the INFINITECH-RA will be built on several concepts that have been introduced by other reference architecture models, including models developed by industrial organizations (e.g., large IT enterprises) and associations (e.g., BDVA), as well as models produced by other projects. Moreover, it considers the structuring principles of BigData and AI applications in digital finance, as these principles are reflected in industrial reference architectures. Nevertheless, the INFINITECH-RA aims at introducing a more flexible approach: Instead of providing a rigorous (but monolithic) structure of BigData/AI applications, it defines these applications as collections of data-driven pipelines. The latter are built based on INFINITECH components. Hence, the INFINITECH-RA provides a number of layered architectural concepts and a rich set of digital building blocks, which enable the development of virtually any BigData or AI application in Digital Finance. The project will offer increased flexibility in defining data-driven applications in the sector, in a way that covers'/subsumes most of the rigorous architectures outlined in earlier paragraphs. [D2.15]

2.2 The INFINITECH Way Innovation

The Innovation is a process that requires to identify and define clear activities from both scientific and industrial domains. The Infinitech project also offer innovative methods and their means for measuring i.e. Infinitech Innovation Readiness Assessment or Innovation Radar, which is construct by Customer Evaluation Levels, IPR readiness Level, Compliance Maturity Level and Infinitech Readiness Level (IRL). In this document we review them and explain its advantages towards catapulting innovation in the financial sector.

The INFINITECH Project (www.infinietch-h2020.eu) has introduced the "INFINITECH Way Foundation" which offers the best practices in FinTech developments for the insurance and financial sectors in Europe.

In a complex and competitive European Ecosystem Infinitech emerge offering efficient and agile methods and introduced an optimized FinTechs-Friendly reference architecture to reduce the problems of dealing with proprietary infrastructures. Infinitech Reference Architecture offers many advantages for further development using the concept of containers and sandbox development and offers a simple-to-follow on-boarding process. A driver of the reference architecture is the innovation and its potential exploitation which should drive the leadership in Europe and beyond borders.

The Innovation is a process that requires to identify and define clear activities from both scientific and industrial domains. The INFINITECH Project also offer innovative methods and their means for measuring i.e. Infinitech Innovation Readiness, Customer Evaluation Levels, IPR readiness Level, Compliance Maturity Level and INFINITECH Readiness Level (IRL). In this document we review them and explain its advantages towards catapulting innovation.

In the other hand Infinitech is using an innovative on-boarding called the Infinitech Way which also act as a methodology outlining containerized solutions and orchestrated DevOps operations and transitioning methods

The Infinietch 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

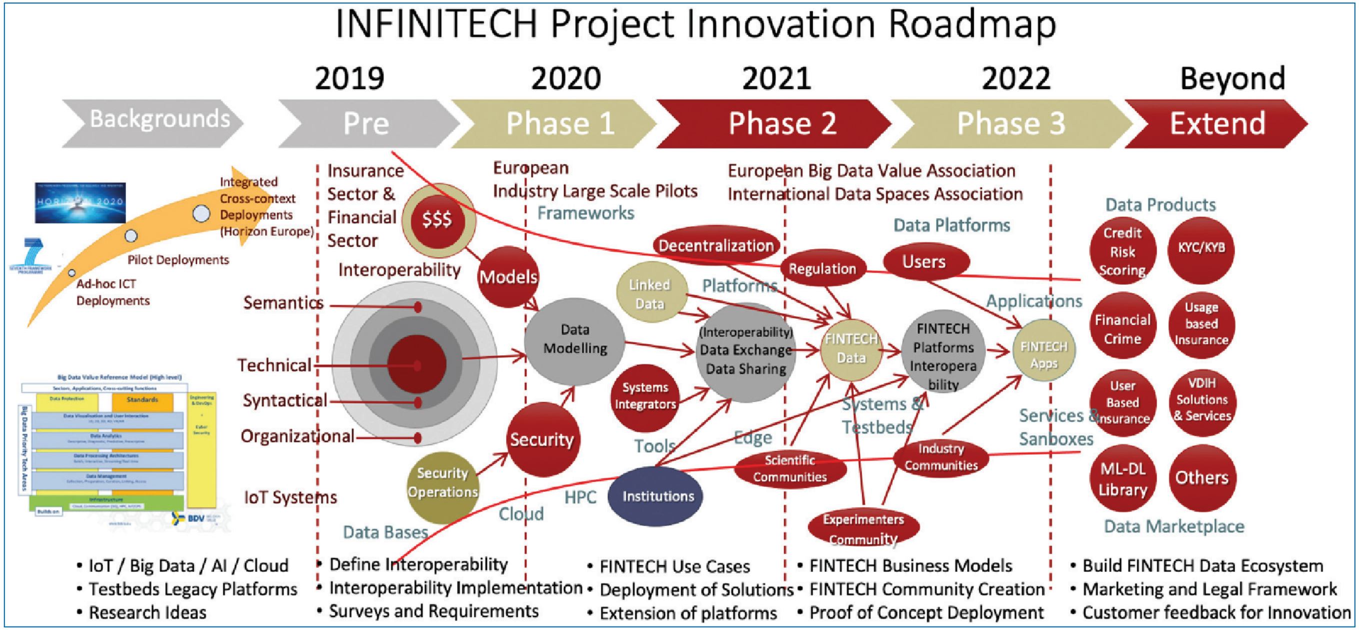

INFINITECH Project Innovation Roadmap
2019
2020
2021
2022
Beyond
Backgrounds
Pre
Phase 1
Phase 2
Phase 3
Extend
HORIZON 2020
Integrated Cross-context Deployments (Horizon Europe)
Pilot Deployments
Ad-hoc ICT Deployments
Big Data Value Reference Model (High level)
Standards
Big Data Priority Tech Areas
BDV
Insurance Sector & Financial Sector
$$$
Interoperability
Semantics
Technical
Syntactical
Organizational
IoT Systems
Data Bases
Security Operations
European Industry Large Scale Pilots
Frameworks
Models
Data Modelling
Security
Cloud
HPC
Decentralization
Linked Data
Platforms
Systems Integrators
Tools
Institutions
(Interoperability) Data Exchange Data Sharing
Edge
Scientific Communities
European Big Data Value Association
International Data Spaces Association
Data Platforms
Regulation
Users
FINTECH Data
FINTECH Platforms Interoperability
Systems & Testbeds
Industry Communities
Experimenters Community
Applications
FINTECH Apps
Services & Sanboxes
Data Products
Credit Risk Scoring
KYC/KYB
Financial Crime
Usage based Insurance
User Based Insurance
VDIH Solutions & Services
ML-DL Library
Others
Data Marketplace
• IoT / Big Data / AI / Cloud
• Testbeds Legacy Platforms
• Research Ideas
• Define Interoperability
• Interoperability Implementation
• Surveys and Requirements
• FINTECH Use Cases
• Deployment of Solutions
• Extension of platforms
• FINTECH Business Models
• FINTECH Community Creation
• Proof of Concept Deployment
• Build FINTECH Data Ecosystem
• Marketing and Legal Framework
• Customer feedback for Innovation

Figure 2.1 INFINITECH innovation roadmap.

research converge 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 complement the activity and bring value to transform ideas an innovate.

2.3 Innovation Radar

Innovation is always difficult to quantity, particularly when multiple and diverse aspects are involved in the innovation process, the Infinitech project is not an exception and thus creative ideas and methods needs to be invented to measure Innovation.

The INFINITECH Pentagon is an adopted kiviat diagram used as the tool to evaluate innovation in the Infinitech project pilots, In the INFINITECH Pentagon, five dimensions are used to evaluate the different relevant areas of a pilot. i,e. Customer, Market, Compliance, Technology, resources and Customer. Each area is assessed and ranked quantitatively, with integers represented in a concentric manner from 1 being the lowest and 5 being the highest.

The Figure 2.2 below shows a graphical representation (called a Kiviat diagram) as example.

After the evaluation of each dimension, each pilot will have its personal Pentagon to validate the current level of Innovation. The methodology to evaluate

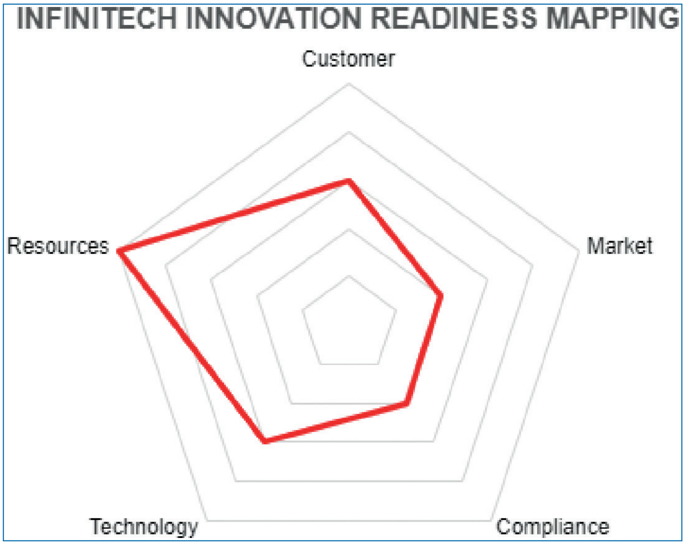


Figure 2.2. INFINITECH pentagon.

each dimension is straightforward – a weighted average combining two different evaluations:

- 50% of the final level is the ranking obtained through stakeholders feedback and evaluation, following the methodology developed and described in T7.8.
- 50% of the final level is the ranking obtained through either quantitative or qualitative analysis and categorization. They will be thoroughly described in the following section.

2.3.1 Customer

It is assessed through three different aspects: the number of Early Adopters, the performed Stakeholders workshops and the formalized interests of possible paying or internal users. Table 2.1 is the categorization and weight of each component.

Table 2.1. Customer evaluation levels.

Categories/Levels	1	2	3	4	5
Early Adopters	0	1–2	3–5	6–9	10+
Stakeholders Workshops	0	1–3	4–8	9–14	15+
Prospects	0–1	2–4	5–9	10–19	20+

Here an example: Pilot#X has 6 Early Adopters, performed 7 Stakeholders Workshops and gathered 18 letters of interest from other stakeholders. The boxes are highlighted in red.

The resulting level of the weighted average is 3,66. Assuming that the level of the feedback assessment resulting from the stakeholders questionnaire (developed in T7.8) is 4, the resulting overall level would be rounded to 4.

2.3.2 Market

The Growth Share Matrix is a fundamental tool to identify the most promising businesses and markets and it is based on the principle that higher and sustainable returns are the core of solid market leadership and that it is possible to obtain

Table 2.2. Pilot 6 evaluation example using customer levels.

Categories/Levels	1	2	3	4	5
Early Adopters	0	1–2	3–5	6–9	10+
Stakeholders Workshops	0	1–3	4–8	9–14	15+
Formalized Interests	0–1	2–4	5–9	10–19	20+

them by analyzing company competitiveness and market attractiveness – through relative market share and growth rate – thus prioritizing investments in markets and business units according to their degree of profitability. Moreover, the matrix underlines the gap that has been generated by the enterprise itself with respect to its competitors, gaining a significant cost advantage on them.

Given that relative market share and growth are the two drivers of the matrix, the latter illustrates specific combinations of them which are displayed in each quadrant. As it is shown by the picture below, markets and business units can be classified as "cash cows", "stars", "question marks" or "dogs":

- Cash Cows are markets/business units with low growth rate and high relative market share which can be leveraged to assign cash to reinvest in more profitable opportunities.
- Stars are characterized by high growth rate and high relative share, being a market/business unit with high future potential.
- Question marks reflect high growth and low share markets/business units which can be further exploited or abandoned according to their probability of becoming stars.
- Pets are low share and low growth markets/business units on which companies should change their position, divest or liquidate.

Growth Share Matrix

The Figure 2.3 depicts the ability of a company to gain a leading position in its market before the growth slows represents a key factor. It is in fact most probable that a product eventually becomes either cash cows or pets. In the Table 2.3

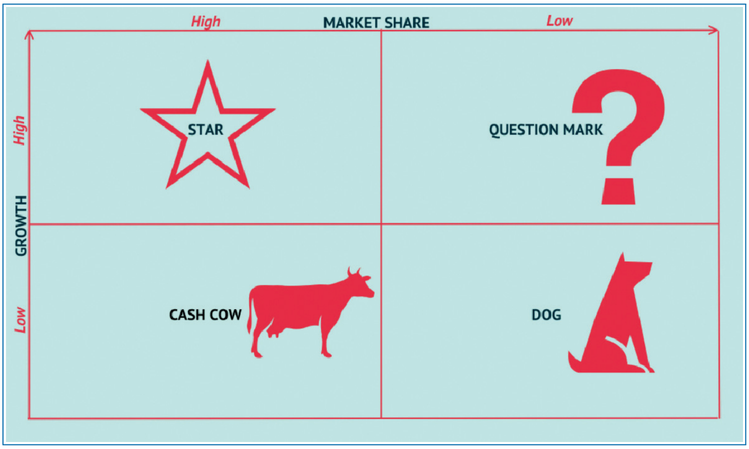
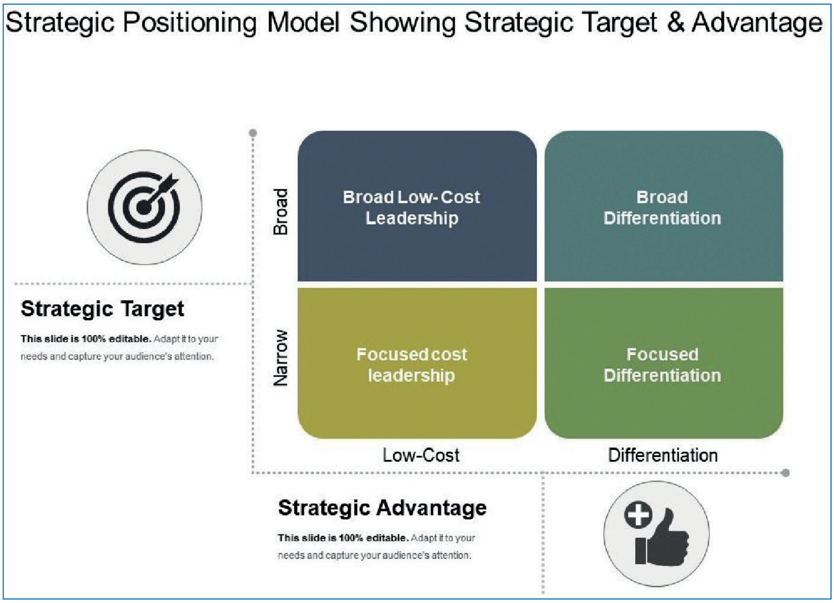


Figure 2.3. INFINITECH growth share matrix example.



Strategic Positioning Model Showing Strategic Target & Advantage

Broad

Narrow

Broad Low-Cost Leadership

Broad Differentiation

Focused cost leadership

Focused Differentiation

Low-Cost

Differentiation

Strategic Target

This slide is 100% editable. Adapt it to your needs and capture your audience's attention.

Strategic Advantage

This slide is 100% editable. Adapt it to your needs and capture your audience's attention.

Figure 2.4. Strategic positioning model.

Table 2.3. Pilot 6 evaluation example using customer levels.

Parameters	1	2	3	4	5
Growth Share Matrix	L;L	L;L	L;H	H;L	H;H
Strategic Positioning Model	B;L	B;L	N;L	B;D	N;D

the first parameter represents growth (L;H) and the second one represents market share (L;H).

Strategic Positioning Model

The Strategic Positioning Model that can be seen in the Figure 2.4 is a tool that helps in the analyses the competitive advantage and the differentiation of a specific solution in respect to competitors.

There are two main variables: the target customers (B – broad and N – narrow) and the perceived value (L – low-cost and D – differentiation) with four possible combinations. INFINITECH is an innovation project and therefore differentiation is its objective. The lowest value is given to the so called Cost Leadership (B;L) on the upper left.

The focused cost leadership (N;L) is considered slightly better; whereas a broad differentiation (B;D) is an improvement and the focused differentiation (N;D) is of the highest value.

2.3.3 Compliance

Nowadays, compliance represents an important tool in the context of competition, as it can provide a competitive advantage. The INFINITECH approach addresses such a key aspect under two perspectives – Intellectual Property Right (IPR) and Compliance Maturity – by leveraging two tools: the IPR Readiness Level and the Compliance Maturity Model.

IPR Readiness Level

The first one assesses the readiness level of IPR by considering the entire process of IPR release, as it is shown by the Table 2.4.

Table 2.4. IPR Levels.

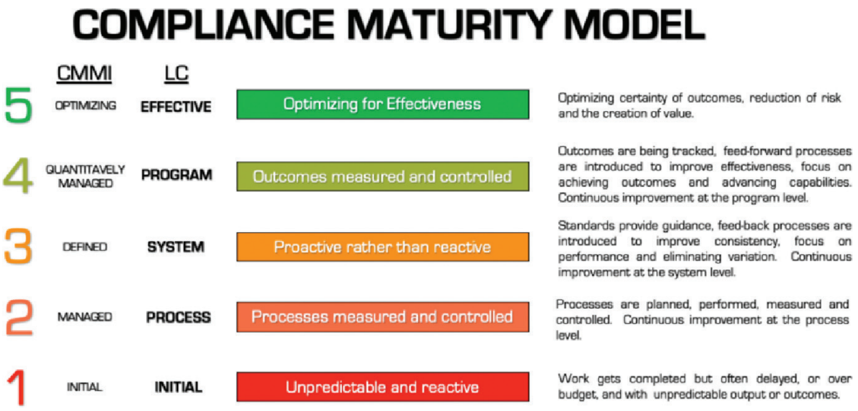
Level	Description
1	Hypothesizing on possible IPR (patentable inventions)
2	Identified specific patentable inventions or other IPR
3	Detailed description of possible patentable inventions. Initial search of the technical field and prior art.
4	Confirmed novelty and patentability; decided on alternative IP protection if not patenting
5	First complete patent application filed, Draft of IPR strategy done.
6	Positive response on patent application; initial assessment of freedom to operate, patent strategy supporting business
7	Patent entry into national phase; other formal IPR registered
8	First patent granted, IPR strategy fully implemented, more complete assessment of freedom to operate
9	Patent granted in relevant countries, strong IPR support for business

Source: https://www.researchgate.net/figure/Intellectual-Property-Readiness-Level-Part-of-TRL-Hasenauer-et-al-Managing_fig4_313063121

Compliance Maturity Model

The second one represents the compliance to internal policies and whether the company working on the solution follows specific procedures. Each pilot should follow specific procedures, have a quality plan and a risk management procedure in place. The guidelines provided by the INFINITECH project can be found in D1.2 and future versions. In order to assess the compliance value, the selected model is the Compliance Maturity Model represented and explained below. It is an adaptation of the widely used CMMI (Capabilities Maturity Model Integration).

Table 2.5. Compliance maturity model.



COMPLIANCE MATURITY MODEL

	CMMI	LC		
5	OPTIMIZING	EFFECTIVE	Optimizing for Effectiveness	Optimizing certainty of outcomes, reduction of risk and the creation of value.
4	QUANTITAVELY MANAGED	PROGRAM	Outcomes measured and controlled	Outcomes are being tracked, feed-forward processes are introduced to improve effectiveness, focus on achieving outcomes and advancing capabilities. Continuous improvement at the program level.
3	DEFINED	SYSTEM	Proactive rather than reactive	Standards provide guidance, feed-back processes are introduced to improve consistency, focus on performance and eliminating variation. Continuous improvement at the system level.
2	MANAGED	PROCESS	Processes measured and controlled	Processes are planned, performed, measured and controlled. Continuous improvement at the process level.
1	INITIAL	INITIAL	Unpredictable and reactive	Work gets completed but often delayed, or over budget, and with unpredictable output or outcomes.

Source: https://www.leancompliance.ca/post/capabilities-maturity-model-for-compliance

Table 2.6. IPR and CMM Levels.

Parameters	1	2	3	4	5
IPR Readiness Level	1	2–3	4–5	6–7	8–9
CM Model	1	2	3	4	5

2.3.4 Technology

Technology represents a key dimension for the Infinitech Innovation Readiness Mapping. Such areas of interest can be split up into two sub-categories: technology readiness level and infinitech readiness level.

TRL Readiness Level

The Technology Readiness Level indicates the technology maturity phase of a technology according to the EU adoption of the NASA TRL developed in the 1970 decade, spanning from basic-principle observations to the actual system proven in an operational environment. The use of TRLs emerge as a guide that enables different background technologist and general public to have a consistent, uniform discussions of technical maturity across different types of technology. TRLs are measured by unity scale of 1 to 9 with 1 being the lowest level and 9 being the most mature technology. The US Department of Defense has used the scale for procurement since the early 2000s. By 2008 the scale was also in use at the European Space Agency (ESA) and in 2010 the European Commission advised EU-funded research and innovation projects to adopt the scale. In 2013, the TRL scale was further canonized by the ISO 16290:2013 standard. TRLs were consequently used in 2014 in the EU Horizon 2020 program. The TRL is determined during a Technology Readiness Assessment (TRA) that examines program concepts, technology requirements, and demonstrated technology capabilities. The second one assesses the actual level of development of a pilot in respect to the Infinitech Way (TBD – we'll have the final version of it in the coming days)

INFINITECH Readiness Level

The INFINITECH Readiness Level (IRL) is a self-assessment method created to evaluate and provide guidance during the adoption process of generic or specific technology assets and/or technology conditions related to the deployment phase of a pilot project. IRL's main objective is to act as a tool for project pilot leaders to help them and guide them in their process to identify and select ready-to-go technologies, in particular IRL also is useful in the on-boarding process of the INFINITECH Way Foundation where technologies that have been developed in advance can be adopted at any stage duration the execution of the project pilots.

IRL is applicable to any program (i.e. INFINITECH Pilot) at any stage that operates or want to make use of the technology assets that are developed or implemented for the financial and insurance sectors, the IRL defines 5 levels scored 1 to 5 where 1 is the lowest level and 5 is the highest level. IRL can be used directly in the INFINITECH Pentagon (i.e. kiviat diagram) as it can be mapped directly within the innovation process/radar facilitating in this way the path towards achieve high levels of Innovation.

On-Boarding Process – the development of ad-hoc technologies is an expensive process and cumbersome if there is not adequate allocated resources (i.e technological and personnel), the re-use of other technologies can facilitate the development and deployment of better applications improving the way teams develop and deliver rapidly and efficiently innovative applications in order to achieve a competitive advantage and operational efficiency.

INFINITECH project has introduced the "INFINITECH Way Foundation" which offers many advantages in the on-boarding process offering speed and agility when optimizing the Reference Architecture and without the problems of dealing with complex, proprietary infrastructure. This on-boarding process outlines adopting containers and orchestration and transitioning implemented solutions to a containerized approach. IRL focus on evaluating the paths is taken in the adoption and deployment of those paths and technologies for an easy adoption and efficient deployment.

Self-Assessment – IRL provides to any project (INFINITECH Pilot) the capacity to self-assess the level of adoption following identified technological characteristics in different technology domain areas i.e. Data Modeling and Data Interoperability, Infrastructure Deployment and Services Platform Adoption, Information Management and Analytics and Intelligent Applications. The technology domain areas are defined by following the requirements from ICT experts into a full-stack implementation and the IRL method allows the self-assessment of the evolution of the pilot project during the phase of adoption or onboarding process. IRL also adds the temporal dimension to an onboarding process which helps to not only self-assess the way the project is being implemented but also the validity of the process for following that path.

IRL Level 1 – Look after the identification and definition of basic conditions to start a pilot with as much as possible re-use of INFINITECH baseline technologies. As a first step other technologies outside INFINITECH Ecosystem may be used but with a clear design perspective for on-boarding Infinitech assets in next iteration.

IRL Level 2 – Ready to go technologies for a prototype building in Client-Server model, a series of demonstrators can be available to test/proof the deployment of basic functionalities, at this level integration is not required but is recommended to have data model integration and common access control tools.

IRL Level 3 – First Deployed Demonstrator with all functionalities in Cloud-based deployment, the use of INFINITECH Way Foundation and Reference Architecture is the core of this level, the use of demonstrators is essential to explain high level use cases and also the use of Data sets following the INFINITECH Data model or other best practices/standards in the particular domain of the pilot demonstration.

IRL Level 4 – Cross Domain services deployed with DevOps ready for easy deployment. Data is required to be shared across different domains, it can be tested with simple applications or integrated in the cross-domain query system. At this level all the components are integrated following DevOps Techniques and the orchestration methods described in the INFINITECH Way Foundation.

IRL Level 5 – Implemented Proof of Concept (PoC) with interoperable cross-domain services, DevOps and Sandboxes deployments are featured to facilitate the inclusion of the assets and solutions in the INFINITECH marketplace.

Table 2.7 summarises the IRL levels providing short descriptions on what is observed and considered for evaluation on each IRL level and its correspondence with technology domains described as parameters and as part of the self-assessment evaluation points.

Table 2.7. Infinitech readiness level (IRL) self-assessment chart.

Parameters	Level 1	Level 2	Level 3	Level 4	Level 5
Data Modeling	Vocabulary Identified	Taxonomy Ready	Data Model Logic & Physical	Validated Data Schema	Data Graph Ready
Data & Interoperability	Data Set Sample Ready	Data Storage Deployed	Query Data Tests Performed	Cross-Domain Query	Data Sharing/ Exchange
Trustworthiness Security & Privacy	Data Protection Methods	Access Control Tools	Identity Management	Platform Access Control	Self-Sovereign identity
Infrastructure	Local Host	Client-Server Mode	Cloud Environment	Docker Ready	Scale Up Tested
Services Platform	Communication Services using Service APIs	Management Services	Continuous Monitoring Tools	Services using Infinitech Orchestration Methods	Infinitech Orchestration and DevOps Compliance
Applications	Use Cases	Prototype	Demonstrator with videos	Online Services	Marketplace Ready

Table 2.8. TRL and IRL levels mapped with innovation pentagon.

Parameters	1	2	3	4	5
TRL	1	2–3	4–5	6–7	8–9
IRL	1	2	3	4	5

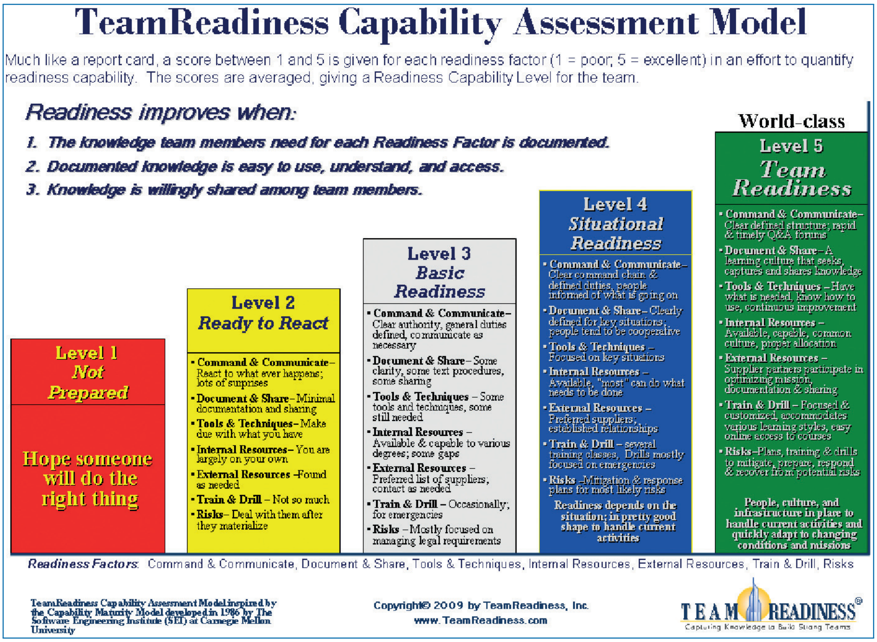
Table 2.8 shows the TRL and IRL mapping with the INFINITECH Innovation Readiness Mapping using the proposed kiviat diagram in Section 3, it is observed how IRL is mapped 1:1 into the innovation assessment radar and intrinsically the evolution of the technology adoption of the pilot in the listed/selected domain areas. Note the listed areas are validated with ICT experts and technical pilot managers in the INFINITECH Project but can also be extended and adequate to particular/specific pilot needs.

2.3.5 Resources

Resources are the fuel that drives the business and can be represented by three key variables: financials; staffing and skills; and organization readiness. The first parameter can be quantitatively assessed by the ROI indicator (the expected or foreseen ROI, in this case) (Table 2.9), further understanding what could be the return of the company on the investments it made: it is a key tool to verify the quality of the decisions made upon managing assets and liabilities.

Table 2.9. Resources levels.

Parameters	1	2	3	4	5
Expected ROI	0–20%	20–40%	40–60%	60–80%	>80%
Team Readiness Capability Level	1	2	3	4	5



TeamReadiness Capability Assessment Model

Much like a report card, a score between 1 and 5 is given for each readiness factor (1 = poor; 5 = excellent) in an effort to quantify readiness capability. The scores are averaged, giving a Readiness Capability Level for the team.

Readiness improves when:

1. *The knowledge team members need for each Readiness Factor is documented.*
2. *Documented knowledge is easy to use, understand, and access.*
3. *Knowledge is willingly shared among team members.*

Level 1 *Not Prepared*

Hope someone will do the right thing

Level 2 *Ready to React*

- **Command & Communicate**– React to what ever happens; lots of surprises
- **Document & Share**– Minimal documentation and sharing
- **Tools & Techniques**– Make due with what you have
- **Internal Resources**– You are largely on your own
- **External Resources** –Found as needed
- **Train & Drill** – Not so much
- **Risks**– Deal with them after they materialize

Level 3 *Basic Readiness*

- **Command & Communicate**– Clear authority, general duties defined, communicate as necessary
- **Document & Share**– Some clarity, some text procedures, some sharing
- **Tools & Techniques** – Some tools and techniques, some still needed
- **Internal Resources** – Available & capable to various degrees; some gaps
- **External Resources** – Preferred list of suppliers; contact as needed
- **Train & Drill** – Occasionally; for emergencies
- **Risks** – Mostly focused on managing legal requirements

Level 4 *Situational Readiness*

- **Command & Communicate**– Clear command chain & defined duties, people informed of what is going on
- **Document & Share**– Clearly defined for key situations; people tend to be cooperative
- **Tools & Techniques** – Focused on key situations
- **Internal Resources** – Available, "most" can do what needs to be done
- **External Resources** – Preferred suppliers; established relationships
- **Train & Drill** – several training classes, Drills mostly focused on emergencies
- **Risks** –Mitigation & response plans for most likely risks

Readiness depends on the situation; in pretty good shape to handle current activities

World-class

Level 5 *Team Readiness*

- **Command & Communicate**– Clear defined structure; rapid & timely Q&A forums
- **Document & Share**– A learning culture that seeks, captures and shares knowledge
- **Tools & Techniques** – Have what is needed, know how to use, continuous improvement
- **Internal Resources** – Available, capable, common culture, proper allocation
- **External Resources** – Supplier partners participate in optimizing mission, documentation & sharing
- **Train & Drill** – Focused & customized, accommodates various learning styles, easy online access to courses
- **Risks**–Plans, training & drills to mitigate, prepare, respond & recover from potential risks

People, culture, and infrastructure in place to handle current activities and quickly adapt to changing conditions and missions

Readiness Factors: Command & Communicate, Document & Share, Tools & Techniques, Internal Resources, External Resources, Train & Drill, Risks

TeamReadiness Capability Assessment Model inspired by the Capability Maturity Model developed in 1986 by The Software Engineering Institute (SEI) at Carnegie Mellon University

Copyright© 2009 by TeamReadiness, Inc.
www.TeamReadiness.com

TEAM READINESS®
Capturing Knowledge to Build Strong Teams

Figure 2.5. Team readiness capability assessment model.

The second and the third are qualitative variables that provide important information on the overall know-how of the company which can be evaluated both in terms of the set of skills the company staff has and whether the company implements and adheres to schedules and detailed documentations, also smoothing internal processes. The self-assessment tool selected for such purpose is the "Team Readiness Capability Assessment Model" (see picture 2.5 above), which quantifies the readiness capability level of team category groups.

There are five levels of Team Readiness which are related to the analysis of the following variables: command and communication, document and share, tools and techniques, internal resources, external resources, training and drilling and metrics. The correspondent levels are the following:

- Lv.1: Not prepared
- Lv.2: Ready to react
- Lv.3: Emergency Readiness

- Lv.4: Situational Readiness
- Lv.5: Team Readiness

2.4 Certification

The world bank defines Financial Sector as a group of resilient, transparent and smooth-functioning financial systems and capital markets that contribute to financial stability, job growth and poverty alleviation. While in the other hand according to Investopedia the Financial Sector is a section of the economy made up of firms and institutions that provide financial services to commercial and retail customers. Any of these two definitions could potentially suit the different stakeholder needs. There is no common agreement and thus not common ways to assess and evaluate innovation in this complex sector. The only common agreement is that this sector comprises a broad range of industries including banks, investment companies, insurance companies, and real estate firms and that in the current business conditions FinTechs are emerging rapidly with very disruptive applications and services best practices for business innovation are required.

In current financial market conditions there is no way to ensure Financial Services are designed with the adequate level of interoperability and that they are deployed following best design practices to compete with the very disruptive FinTech Markets. However there are good indicators coming from the INFINITECH Innovation Readiness, INFINITECH On-boarding Process and INFINITECH Readiness Level approaches that may constitute the first step to create the certification program that is required to ensure that best practices are followed, standards are used and the created, implemented, deployed and tested technologies within the INFINITECH Marketplace are used.

Figure 2.6 shows the potential certification label that will be provided to pilots that demonstrate (following the certification programme) to be part of the INFINITECH ecosystem but beyond that to prove that all their technologies follows the INFINITECH way and that the provided applications at the plot level are fully compliant with innovative levels.

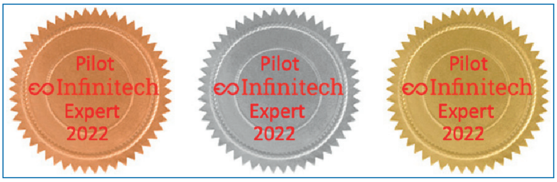


Figure 2.6. INFINITECH innovation-certification programme.

2.5 Infinitech Terminology

The common vocabularies in the insurance and financial sectors are evolving in the same way society and economy, INFINITECH project take an step ahead and identify and define terminology that is required for innovation activities, from simple concept like assets and case studies to a more elaborate terms like integration and marketplace, terms that are emerging as trend activities in software systems. Table 2.10 summarises the Infinitech terminology which can help for better understanding at INFINITECH project level but also in other activities in FinTechs.

Table 2.10. INFINITECH adopted terminology.

Concept/Term	Definition
Asset	A technology that was implemented or extended in the context of INFINITECH project.
Case Studies	An Example (Idea, Slides, Diagram, Video, Demo) How an INFINITECH Technology within a pilot works.
Demonstrator/ Demo	An implemented Example in how a Pilot potentially could implement the Case Studies
PoC	The implemented concept(s) or use cases for a Pilot progressing a Demo, usually case studies are the main concept of a pilot and use cases are short experiences that can be proven describing the case studies as a Demo
Use Cases	An implemented experiences that can be proven be part of a larger case Study idea and with the objective to provide insight towards integrated PoC.
Exploitation	The potential of a Case Study for being implemented and commercialised as part of a INFINITECH PoC implementation.
Integration	The PoC showing in how to use one or varios Infinitech technologies (i.e. Integrated Demo), note that Demos that do not use Infinitech technologice remains in the category of Demo
Deployment	A Demo or PoC installed/instantiated in an INFINITECH Server or at Stakeholder Infrastructure
Service	The networking conditions or software tools for implementing a PoC following Case Studies
Operation	A Demo running in Infinitech Server or Stakeholder Infrastructure
Application	A group of software programs performing functions towards the Use Cases in the PoC.
Marketplace	Infinitech server or stakeholder Infrastructure showing assets i.e. data, platforms, components, pilot results, etc

DOI: 10.1561/9781638282297.ch3

Chapter 3

Reference Architectures Analysis

3.1 Reference Architecture Analysis

Reference Architectures 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:

- 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 organisation may accumulate technical risks and end up with a complex and non-optimal implementation architecture.

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 book aims to partially cover the gap.

3.2 Methodology

The methodology for the development of the INFINITECH-RA is based on the following overlapping phases:

- Phase 1 – Review and Analysis (M2-M7): This phase was focused on the analysis of relevant technical and scientific information, ensuring that the INFINITECH-RA considers recent developments in BigData architectures in general and for the finance sector in particular. It analysed the BDVA RA, along with BigData architectures for digital finance applications that have been introduced and are widely used in the industry. In this context, BigData/AI environments and tools have been reviewed as well. Likewise, this phase considered the stakeholders requirements (from T2.1 of the project), as well as regulatory compliance requirements and the specifications of INFINITECH technologies reflected in D2.5 and D2.7 respectively. The overarching goal of this phase was to make sure that the design of the INFINITECH-RA aligns with key requirements of BigData applications in digital finance, as well as with the evolution of the state of the art in BigData in finance.
- Phase 2 – Architecture Design (M6-M12): This phase produced the initial design of the INFINITECH-RA, using the 4 + 1 view methodology [Kruchten95] for specifying software architectures. The selection of this proven and well-known methodology was motivated from the fact that INFINITECH pilot systems are essentially software systems. Hence, a methodology for specifying software architectures was applicable. In-line with the 4 + 1 architecture the INFINITECH-RA is described as a collection of complementary views (i.e. process, development, logical, physical) that represent dynamic behaviors and static behavior of the systems, as well as relevant implementation aspects. Likewise, 4 + 1 signifies the need for confronting the architecture against a number of scenarios like the INFINITECH pilots. As part of the second phase of the INFINITECH development, a set of main views for the INFINITECH-RA were specified and described, as presented in subsequent paragraphs. Note that the phase extends beyond the delivery of the first version of the INFINITECH-RA

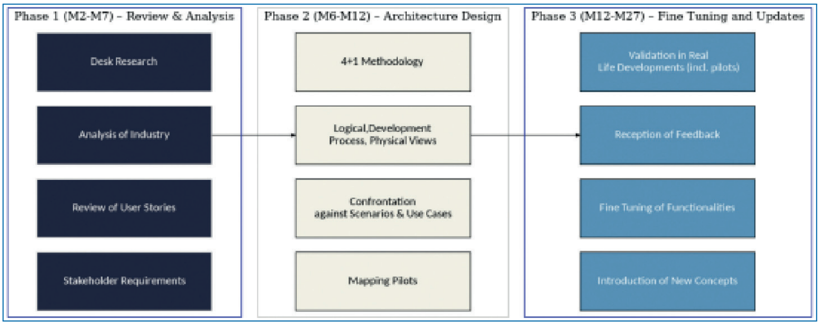

Phase 1 (M2-M7) - Review & Analysis
Desk Research
Analysis of Industry
Review of User Stories
Stakeholder Requirements
Phase 2 (M6-M12) - Architecture Design
4+1 Methodology
Logical,Development Process, Physical Views
Confrontation against Scenarios & Use Cases
Mapping Pilots
Phase 3 (M12-M27) - Fine Tuning and Updates
Validation in Real Life Developments (incl. pilots)
Reception of Feedback
Fine Tuning of Functionalities
Introduction of New Concepts

Figure 3.1. High-level overview of the project's phased approach to INFINITECH-RA development.

(i.e. to M12 of the project), as several of the presented views will be revised in the coming months.

- Phase 3 – Fine Tuning and Updates (M12-M27): The third development phase of the INFINITECH-RA started after the submission of D2.13. Phase 3 aims at revising and fine-tuning the INFINITECH-RA based on its actual use in other work packages, including the sandboxes and pilot development work packages. It will receive and exploit stakeholders' feedback from the practical use of the INFINITECH-RA for developing BigData/AI systems for the finance sector. The feedback will be exploited towards improving aspects of the INFINITECH-RA, but also towards validating the architectural concepts introduced in practice. This Phase will lead to the final version of the INFINITECH-RA as part of deliverable D2.15 i.e. the third and last version/release of this deliverable.

The Figure 3.1 illustrates the three development phases of the INFINITE CH_RA and the main activities that they comprise.

3.3 General AI/BigData Challenges

3.3.1 Big Data

Over the past few years and in the context of big data, new challenges have emerged in data analytics, turning the attention to methodologies that raise the abstraction level and facilitate the convergence rate of algorithms. To this end, several big data frameworks have been developed by researchers and engineers and respective algorithms have been compiled, mainly being domainspecific. However, the development of scalable and distributed analytics solutions for extreme-scale analytics in the finance/insurance sector remains quite a complicated process given the diversity in datasets and data sources.

Data sources can be of different type (structural, semi-structural or completely unstructural), in various formats and data accessibility options. Due to the abovementioned factors, , modern enterprises rely on a variety of different and heterogeneous data management systems in order to handle with this data diversity: relational datastores that are used to store structural data compliant with an entity-relational model where the ensurance of transactional semantics and data consistency is of major concern, document-based data management systems to store semi-structural data, keyvalue stores that are considered efficient to store data coming from IoT sensors or logging information (i.e. when navigating among different web pages in the network, or simply logging the details of a finance transaction that took place) and even the use of HDFS data lakes is now considered prominent and facilitates the analysis of unstructural data that can be available. However, the analysis over a superset of the available data management system of an organization is not a trivial task. Joining data across datasets is very cost demanding and difficult to be implemented efficiently in the application or data processing level. To overcome this barrier, various analytical frameworks have been proposed that provide polyglot capabilities and abstract this problem from data scientists, making the data management process transparent from their perspective.

Accessing heterogeneous data sources (a concept often addressed by data integration systems or multidatabases [Özsu11], [Tomasic98]) is a problem that has been widely studied in the literature and with the recent emerge of cloud database and big data processing, it has been evolved towards polystore systems. Early implementations [Minpeng11], [Ong14], [Simitsis12] relied on a single common model that the target datastores had to transform their schema to. A further improved presented by the polystore BigDAWG [Duggan15], [Gadepally16] which defines islands of information, which makes use of a single data model and language. Nowadays, Spark SQL [Armbrust15] exploits its advantage for massive parallel processing over a federation of different and heterogeneous stores. It defines the notion of dataframe and provides different connectors over a variety of supported datastores. By providing a single interface and a common query language, it pushes down the query execution on the target databases, when this is possible, and retrieves data into those dataframes that are being used for further processing. That requires however the retrieval of the data in the data analysis layer, which can be memory greedy. Similarly, Presto allows for massive parallel processing where a coordinator orchestrates the query execution on several workers which make use of corresponding connectors, all of them implementing a common interface in order to hide the details of the query execution on the target database. Towards the same direction, Apache Drill and Impala maintain the notion of the data connectors that are able to retrieve data from external and heterogeneous datastores, and transform it to their own intermediate format and model that can be used for data processing in the upper layer of the framework. All those approaches provide polyglot capabilities, however, they are considered as additional frameworks that require the retrieval of the dataset in memory, and exploit their abilities for massive parallelism to be able to scale out adequately in order to deal with these requirements. The challenge lies on exploring the attributes and statistics of each individual dataset that is stored in each store, optimize the query execution and retrieve only the minimum amount of data that is needed in each query operator, thus minimizing the need for memory consumption and data traffic in the network.

Besides the analytics level, running analytics over stale data is not ideal. For example, consider the case of a recommender system that aims to predict the behavioural patterns of a user, her preferences or dislikes, and provide personalised recommendations of relevant items. In a real-world scenario, user preferences change frequently, and new data continuously arrive in a real-time manner. A recommender system should, if possible, adapt to these changes as they happen, modifying its model to always reflect the current status, while requiring a single pass through the data. Furthermore, training new versions of the model in a continuously growing dataset is computationally inefficient, leading also to unnecessarily amplified infrastructure costs. Hence, data management approaches are required both to account for new datasets and to support these incremental analytics.

Apart from the scenarios where systems and analytical algorithms should continuously take into account the overgrowing datasets in order to update the results of their analysis, they will also need to analyze real-time data in order to respond to events and create alerts and notifications as they events occur. An example of such a scenario is fraud detection, where users must be notified when the fraud transaction occurs. However, dealing with data related to financial transaction raises a lot of concerns regarding data consistency and isolation. This is the reason why traditional relational database management systems are used, which ensure transactional (from a database theory perspective) semantics and ensure that a series of actions will be executed all or none (Atomicity), the data will be consistent (Consistency) and that they will not be stored breaking the rules that have been imposed by the database administrator (the balance of a bank account cannot be negative), parallel actions can be executed concurrently, but the result of their modifications will be likewise each action was executed in order (Isolation) and finally, that data will be preserved in terms of a system failure and all information can be recovered (Durability). Those four attributes consist of the ACID properties that traditional database management systems ensure and are of major importance in the finance sector. These databases are ideal for handling operational workloads, and they rely on the use of locks on data items that are being modified when a transaction access a data table. However, in order to perform a fraud detection, the system needs to read a variety of data by scanning the whole data table. In other words, to perform an analytical operation.

Operational and analytical operations are compatible and one blocks the other, as the existence of a lock will prevent the latter from being executed. Likewise, a long-running analytical operation will pose read locks in the whole data table, thus blocking all operational workloads from being executed: the database will not be able to serve customer needs as long as the long-running operation takes place. Due to this, modern enterprises take a snapshot of the operational database and send it to a data warehouse, thus separating those two workloads: The relational datastore will serve operational workloads by ensuring the ACID-ity, while the data warehouse will serve analytical query processing. The drawback of this approach is that detection of a fraud transaction will take place the following day, as moving data from one store to the other is cost-expensive operation and relies on heavy ETLs that are being performed during the night. Even exploiting modern approaches and apply microbatching in the data movement process (i.e. the modifications of a transaction are being sent to a data queue, and a worker periodically gets data from the queue and sends them to the analytical warehouse) this allows for near real-time analytical processing, and cannot detect events that happen on real-time. Due to this, the analytics over operational data is still of great challenge.

3.3.2 Data Pipelines

An additional aspect and challenge refers to data management that needs to be enhanced in order to both facilitate the needs at extreme-scale (in terms of efficiently for data throughput and access) and to address the challenge of various federated and distributed systems that hold and corresponding datasets. What is more, one needs to consider the changes on the underlying systems (and datasets) and thus the required level of dynamicity, raising the challenge for approaches that perform data pipelining from heterogeneous distributed systems towards analytics frameworks, while being adaptive to the aforementioned changes.

Tightly-coupled multistore systems trade autonomy for performance, typically in a shared-nothing cluster, to integrate structured (RDBMS) and HDFS data. Polybase [Minukhin18] is a feature of Microsoft SQL Server Parallel Data Warehouse to access HDFS data using SQL. It allows HDFS data to be referenced through external PDW tables and joined with native PDW tables using SQL queries. Hybrid systems are similar to tightly-coupled systems, e.g. integrating HDFS and RDBMS in a sharednothing cluster, except that the HDFS data is accessed through a data processing framework like MapReduce. For example, QoX [Xu18] uses a dataflow approach for optimizing queries over relational (RDBMS and ETL) data and unstructured (HDFS) data, with a black box approach for cost modelling.

Moreover, when dealing with analytics performed on a variety of data sources, it is no less important to focus on the data aspects. Data-intensive distributed frameworks such as such as Apache Spark and Apache Drill can access multiple data stores using a unified API such as SQL. However, applications running on these frameworks have direct access to specific data stores and as a result to specific datasets, while frameworks such as Apache Ranger offer standardized access authorization to data stores, but only for a limited set of supported data stores and with limited policies.

What is required refers to an approach for data management that minimizes the data pipelining process and would enable a hybrid management of data, both for analytical and for transactional workloads. The latter would enable analytics to account for the different datasets made available as they are ingested in the data stores (mainly through the respective transactions).

3.4 Specific Challenges for the Finance Sector

3.4.1 Siloed Data and Business Operations

One of the most prominent challenges faced by banks and financial organizations is the fragmentation of data across different data sources such as databases, data lakes, transactional systems (e.g., ebanking) and OLAP (On Line Analytical Processing) systems (e.g., Customer Data Warehouses). This is the reason why financial organizations are creating BigData architectures that provide the means for consolidating diverse data sources. As a prominent example, the Bank of England has recently established a "One Bank Data Architecture" based on centralized data management platform . This platform facilitates the BigData Analytics tasks of the bank, as it permits analytics over significantly on larger datasets.

Note that the need for reducing data fragmentation has been also underlined by financial institutions following the financial crisis of 2008, where several financial organizations had no easy way to perform integrated risk assessments as different exposures (e.g., on subprime loans or ETFs (Exchange-Traded Fund)) were siloed across different systems. INFINITECH-RA must therefore provide the means for reducing data fragmentation and taking advantage of previous siloed data in integrated BigData Analytics and ML tasks.

3.4.2 Real Time Performance Requirements

Real Time Computing is when an IT system must respond to changes according to definite time constraints, usually on the order of milliseconds or seconds. In the realm of financial and insurance sectors, real time constraints apply where a response must be given to provide services to users or organizations and are in the order of seconds or more. Examples range from banking application to cybersecurity. Most of real-world financial applications are NOT real time as other industrial automation (plant control) and are usually solved putting more computing resources (cpu/gpu power, memory,...) at the problem. However, in the case of ML/DL and BigData, algorithms can take significant amount of time and become useless in practical cases (responses arrive too late to be used). In those cases, a quantitative assessment of the computing time of algorithms is needed to configure resources to provide acceptable time.

3.4.3 Mobility

The digital transformation of financial institutions includes a transition to mobile-first banking [Bons12]. This refers to the interaction of customer and financial organizations through mobile channels as part of mobile banking. INIFINITECH must support mobile channels when visualizing BigData, AI and IoT applications for digital finance, but also when collecting and processing user's/customer's input.

3.4.4 Omni-Channel Banking – Multiple Channels Management

One of the main trends in banking and finance is the transition from conventional multi-channel banking to omni-channel banking [Komulainen18]. The latter refers to seamless and consistent interactions between customers and financial organizations across multiple channels. Hence, omnichannel banking/finance focuses on integrated customer interactions that comprise multiple transactions, rather than individual financial transactions. The INFINITECH-RA should provide the means for supporting omni-channel interactions through creating unified customer views and managing interactions across different channels based on integrated/consolidated information about the customer. Note that BigData analytics is the cornerstone of omni-channel banking as it enables the creation of unified views of the customers and the execution of analytical functions (including ML) that track, predict and anticipate customer behaviors.

3.4.5 Automation

Data intense applications, also in the banking and insurance sectors, are realized by specialized IT and RDBMS administrators. Recently, more and more data scientists and business analysts are involved in the development of such applications at a significant cost. The INFINITECH-RA should provide the means for orchestrating data intense application and management through easily creating workflows and data pipelines.

3.4.6 Transparency

During the last couple of years, financial organizations and customers of digital finance services raise the issue of transparency in the operation of BigData systems as a key prerequisite for the wider adoption and use of BigData analytics systems in financial, sector use cases. This is particularly important for use cases involving the deployment and use of AI/ML systems that operate as blackboxes and are hardly understandable by finance sector stakeholders. Hence, a key requirement for AI/ ML financial sector applications is to be able to explain their outcomes. For example, a recent paper by bank of England [Bracke19] illustrates the importance of providing explainable and transparent credit risk decisions.

Overall, INFINITECH shall provide support for transparency in AI/ML workflows based on the use of Explainable Artificial Intelligence (XAI) techniques such as LIME (Local Interpretable Model-agnostic Explanations) [Ribeiro16], which develops/finds a local model around the prediction that is interpretable. XAI techniques should be supported as part of the ML techniques of the project, but also as an add-on to them as in several cases they will be used to interpret the outcomes of other ML/ AI techniques (e.g., some classifiers).

3.5 Reference Architecture for Big Data/IoT/AI in Finance/Insurance

INFINITECH introduces and validate a 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-Online Analytical Processing, Data Lakes, SQL databases, NoSQL databases, alternative data sources), efficient real-time predictive analytics, multi-channel/Omnichannel 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.

3.6 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. [D2.13]

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. [D2.13]

High-Performance Computing (HPC) can be distributed at nodes within the platform supporting a high degree of scalability. 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 will be identified where existing technologies are available while other components will be designed, implemented, and integrated according to different assigned tasks. [D2.13]

The validity of the RA has been proved by a mapping the workflows of the pilots of the projects, ultimately proving the conceptual approach of the INFINITECH RA. [D2.13]

The RA constitutes a living solution constantly verified during the continuous project development and in particular 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. [D2.15]

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 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. 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 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.

The INFINITECH Reference Architecture (IRA) (shown in Figure 3.2 and explained in Deliverable D2.13), defines a set of layers and allows flexible workflows out of data processing modules, called components that performs specific data transformations. A workflow usually consumes one set of data (sources) to produce another set of data (destinations). A component can be identified by the Input/Output interfaces and the functionality (transformation) provided at the edge. In that respect, a data component can be considered as a black box that can be replaced by any technological implementation that performs at the edge the same transformation. [D2.13]

The Figure 2.1 depicts the different reference layers of the IRA along with examples of data components. Scope of the section is to identify existing and desired data components that will support the design, the development and deployment of the Pilots' sandboxes. [D2.13]

The "black-boxes" follow predefined colors in order to map components in the different layers. An INFINITECH solution (sandbox) can be built organizing components in a workflow (sometimes referred to data pipeline) to accomplish a complex transformation from one set of data sources to another set of data solving the business case. Components should therefore be interoperable, with clear interfaces and perform a clear function over the data. The rest of this section is a first attempt to identify the existing and to be developed components of the project. Further progress in the project will provide the needed details for implementation and deployment.

The first step is the ingestion of data from the relevant data sources, typically stored in operational databases. This data ingestion is massive and takes many hours, a significant amount of time. The challenge here is to be able to accelerate data ingestion without sacrificing the querying speed that becomes crucial for the next steps.

The second step is data protection that enables coordinating the invocation of components that implement privacy, security, or data protection techniques as well as other external services in order to provide a suitable privacy, Security and data protection level specified by a secure service provider compliance to regulations. Further, the data Protection Orchestrator coordinates several privacy, security and data protection components and services to ensure that the successive use of the data that have been protected can be processed or stored preserving their privacy and Security. It also allows the removal of the protection of the results (if required) before delivering them to the end user.

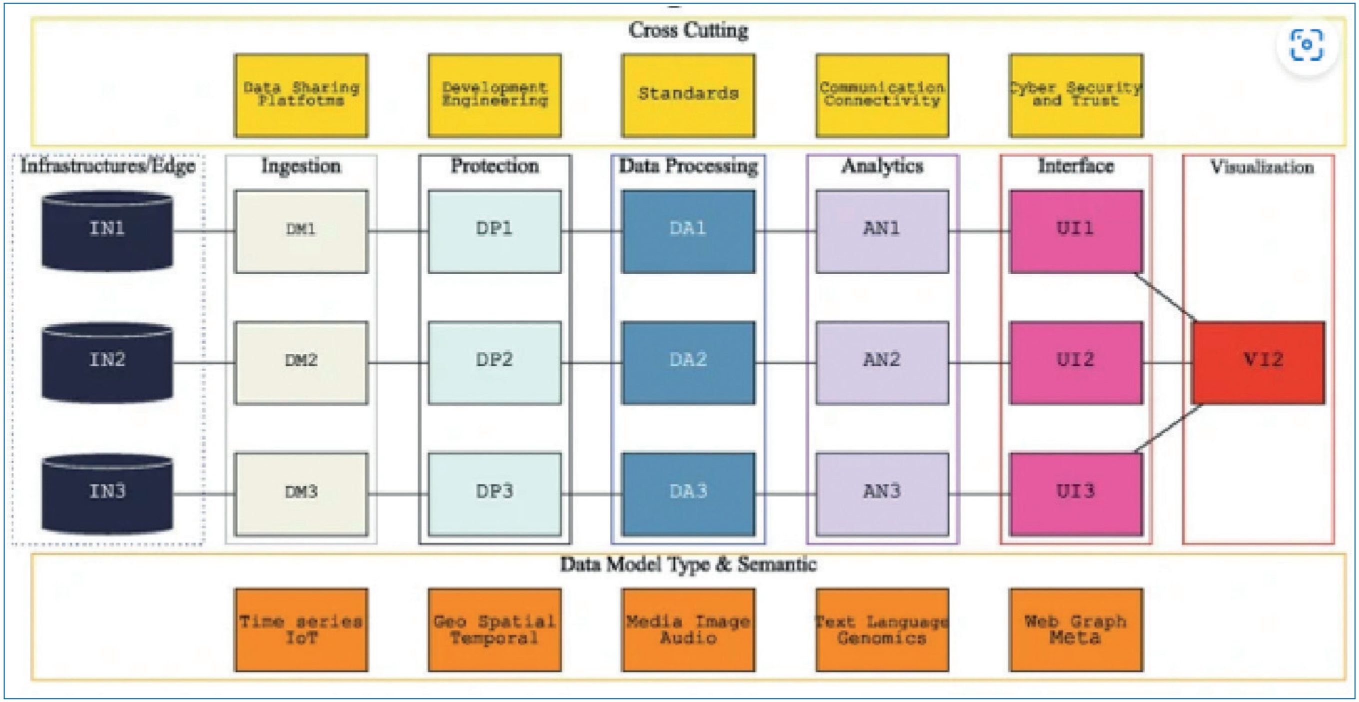

Cross Cutting
Data Sharing Platfotms
Development Engineering
Standards
Communication Connectivity
Cyber Security and Trust
Infrastructures/Edge
Ingestion
Protection
Data Processing
Analytics
Interface
Visualization
IN1
IN2
IN3
DM1
DM2
DM3
DP1
DP2
DP3
DA1
DA2
DA3
AN1
AN2
AN3
UI1
UI2
UI3
VI2
Data Model Type & Semantic
Time series IoT
Geo Spatial Temporal
Media Image Audio
Text Language Genomics
Web Graph Meta

Figure 3.2 Logical view of INFINITECH-RA and mapping to BDVA RM.

The next step is data processing. The development of ML uses cases for digital finance entails a series of data processing steps, which can be structured in end-to-end data management pipelines. INFINITECH-RA specifies building blocks for processing large datasets from their ingestion to the ultimate visualization of ML results. This approach is in-line with the operation of most ML platforms and tools (e.g., KNIME, Auto-SKLearn, MLBox) which facilitate the development and integration of end-to-end pipelines. INFINITECH-RA builds on popular ML concepts, as well as on best practices introduced by relevant Reference Architecture Models for the finance sector that have been introduced in the industry.

INFINITECH's Data Analytics Platforms are products developed to leverage the promptness and accuracy of the suspicious events recognition while reducing costs and optimizing procedural efforts.

The interface group contains projects that provide UI functionalities of general purpose. This might include some user interfaces, visualization graphs that can be reused by pilots to show the results of analytical tools etc. It is important to highlight here that this group does not include front-end implementations that are specific to the pilots, rather than implementations that are generic and can be reused by each pilot.

The data visualization is a layer that deals with visualization of data.

In summary The INFINITECH-RA defines layers to logically group components. The identified layers are: [D2.13]

- **Data Sources:** At the infrastructure level there are the source of data (database management systems, data lakes holding non-structural data, etc).
- **Ingestion:** A layer of data management usually associated with data import, semantic annotation, and filtering from data sources.
- **Security:** A layer for management of the clearance of data for security, anonymization, cleaning of data before any further storing or elaboration
- **Management:** A layer responsible for the data management aspects, including the persistent storage in the central repository and the data processing enabling advanced functionalities such as Hybrid Transactional and Analytical Processing (HTAP), polyglot capabilities, etc
- **Analytics:** A layer for AI/ML components.
- **Interface:** A layer for the definition data to be produced for user interfaces.
- **Cross Cutting:** A layer with service components that provides functionalities orthogonal to the data flows (e.g. Authentication, Authorization, …).
- **Data Model:** A cross cutting layer for modelling and semantics of data in the data flow.
- **Presentation/Visualization:** A layer usually associated with the presentation applications (e.g., desktop, mobile apps, dashboards)

In the end one should note that there are horizontal and vertical concerns.

- **Horizontal concerns** cover specific aspects along the data processing chain, starting with data collection and ingestion, and extending to data visualisation. The horizontal concerns do not imply a layered architecture. As an example, data visualisation may be applied directly to collected data (the data management aspect) without the need for data processing and analytics.
- **Vertical concerns** address cross-cutting issues, which may affect all the horizontal concerns. In addition, vertical concerns may also involve non-technical aspects.

DOI: 10.1561/9781638282297.ch4

Chapter 4

INFINITECH Data Pack

4.1 INFINITECH Data Pack

The INFINITECH Data Pack is the set of files, schemas, and metadata model diagrams (Graphs) that represent the way the INFINITECH data is organised and structured. It contains the metadata in .ttl format and also contains the metadata in two different formats, .json-ld and .owl to ensure the Data Pack is accessible to different communities. [D4.3]

The INFINITECH Graph Data Model is the documentation that describes in detail all the taxonomies and vocabularies from INFINITECH Core, FIBO (https://edmcouncil.org), FIGI (https://www.openfigi.com) and LKIF (https://github.com/RinkeHoekstra/lkif-core) domains used in INFINITECH and that describes and represent all the relationships between them to build the Data Representation of the INFINITECH Graph Data Model. [D4.3]

The INFINITECH Core ontology compile the common vocabularies that are used across different fintech's and finance domains in the INFINITECH Project. The objective of the INFINITECH Core Ontology is to summarize the common terms or vocabularies and establish their relationship based on similarities or equivalences. [D4.3]

This ontology is the compilation of vocabularies and taxonomies that are common to the INFINITECH project and that are used in fintech's and finance domains. The INFINITECH core ontology main purpose is to enable data interoperability and data exchange by identifying the similar vocabularies that are used across different domains but because they pertain to different domains, the similar terms are not related. INFINITECH core establishes and define the relationships between them, facilitating the identification and definition of data exchange and data sharing. [D4.3]

4.1.1 FIBO, Financial Industry Business Ontology

FIBO is a business conceptual ontology standard providing a description of the structure and contractual obligations of financial instruments, legal entities, market data and financial processes. The primary application of the business conceptual ontology is for data harmonization and for the unambiguous sharing of meaning across data repositories. This common language (or Rosetta stone) for the financial industry supports business process automation and facilitates risk analysis. [https://github.com/edmcouncil/fibo]

Figure 4.1 depicts the components of FIBO including foundations, business entities, securities , and financial business and commerce.

4.1.2 FIGI Financial Instrument Global Identifier

The Financial Instrument Global Identifier is an open standard, unique identifier of financial instruments that can be assigned to instruments including common

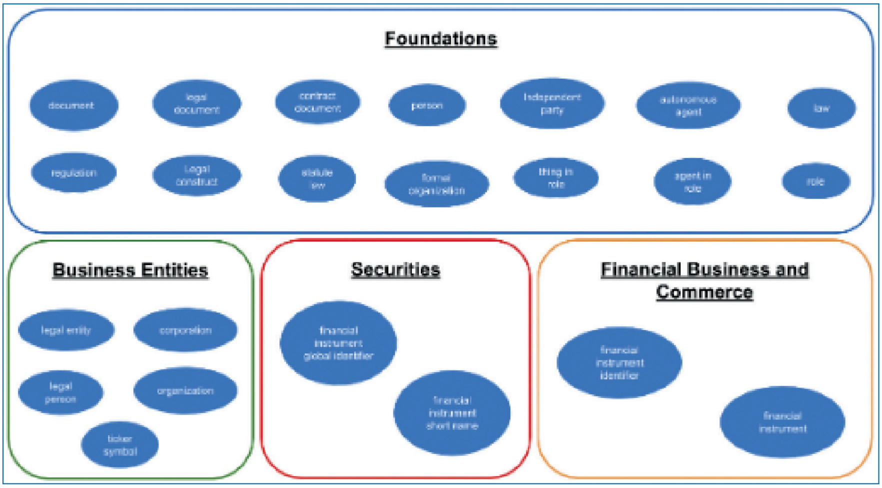

Foundations
document
legal document
contract document
person
Independent party
autonomous agent
law
regulation
Legal construct
statute law
formal organization
thing in role
agent in role
role
Business Entities
legal entity
corporation
legal person
organization
ticker symbol
Securities
financial instrument global identifier
financial instrument short name
Financial Business and Commerce
financial instrument identifier
financial instrument

Figure 4.1. Financial industry business ontology (FIBO).

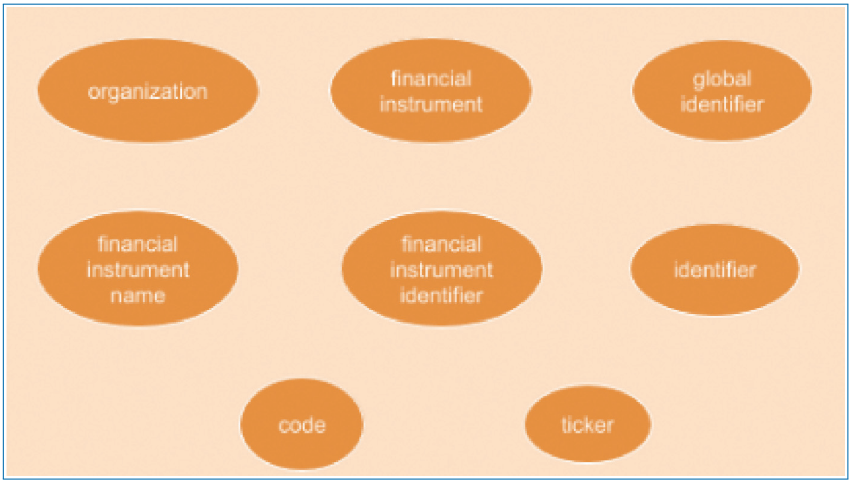


Figure 4.2. Financial instrument global identifier – FIGI.

stock, options, derivatives, futures, corporate and government bonds, municipals, currencies, and mortgage products. [https://www.openfigi.com/api]

Figure 4.2 illustrates the component of FIGI including organizations, financial instruments with its name and identifier, global identifier, identifier, code, and ticker.

4.1.3 LKIF, the Legal Knowledge Interchange Format

Legal Knowledge Interchange Format (LKIF) models legal rules of the kind found in legislation and regulations. LKIF is an Upper and Core ontology from the the ESTRELLA project. The OWL files are available on GitHub. [https://github.com/RinkeHoekstra/lkif-core] Figure 4.3 shows the components of LKIF including norm, action, expression, legal action, and role.

4.1.4 INFINITECH Core

The INFINITECH Core is the baseline of vocabularies and taxonomies used in the INFINITECH Project and that summarizes the different concepts that are overlapping the different financial and insurance tech. INFINITECH Core indicates whether the component is reusable and part of the core INFINITECH infrastructure or a pilot specific component. The INFINITECH Core is a top-bottom approach in order to identify all the vocabularies that are common cross ontologies and define the reference INFINITECH Graph Data Model implementation. [D4.4 to D4.6]

Rather than thinking in extending the reference ontologies and/or identifying new things/concepts and relations it is a better approach to create an INFINITECH

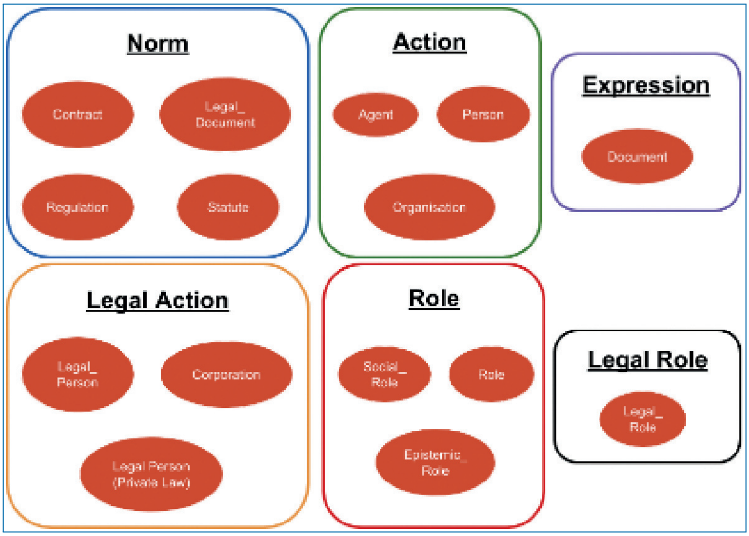

Norm
Contract
Legal_ Document
Regulation
Statute
Action
Agent
Person
Organisation
Expression
Document
Legal Action
Legal_ Person
Corporation
Legal Person (Private Law)
Role
Social_ Role
Role
Epistemic_ Role
Legal Role
Legal_ Role

Figure 4.3. Legal knowledge interchange format – LKIF.

core ontology that grounds on top of the three reference ontologies. To do that, it is necessary to identify without any reasonable doubt all these commonalities while connecting them in one main model, the so-called INFINITECH core model. Figure 2.5 shows this process and the connections between FIBI, FIGI, and LKIF. The picture summarizes the relationship between INFINITECH core component based on subclass, property, and alignments. [D4.4. to D4.6]

Figure 4.4 represents the conceptual alignments between FIBO, LKIF and FIGI. All the three ontologies have common concepts between them. For example, organization concept is defined in all the three ontologies. We have identified the common concepts and defined relationships between the common concepts. In Figure 2.5, you can see that “equivalentClass” relationship is defined between the concept representing an organization in FIBO, FIGI and LKIF. In the case of document concept, “subClassOf” relationship between document concept from FIBO and Document concept from LKIF is defined.

The bottom part of Figure 2.5 shows the diagram legend. Classes are represented as oval shapes, while properties (relationships) between classes are represented using solid lines with filled arrowheads on one side of it to show the direction of the relationship. The property (relationship) is shown in the rectangular box attached to these lines. The “subClassOf” relationship is represented using a dotted line with hollow arrowhead on the side of super class. The color of a relationship represents the origin of the relationship, i.e. green color relationships represent alignments defined by us while black colored relationships come from the respective ontology. [D4.1]

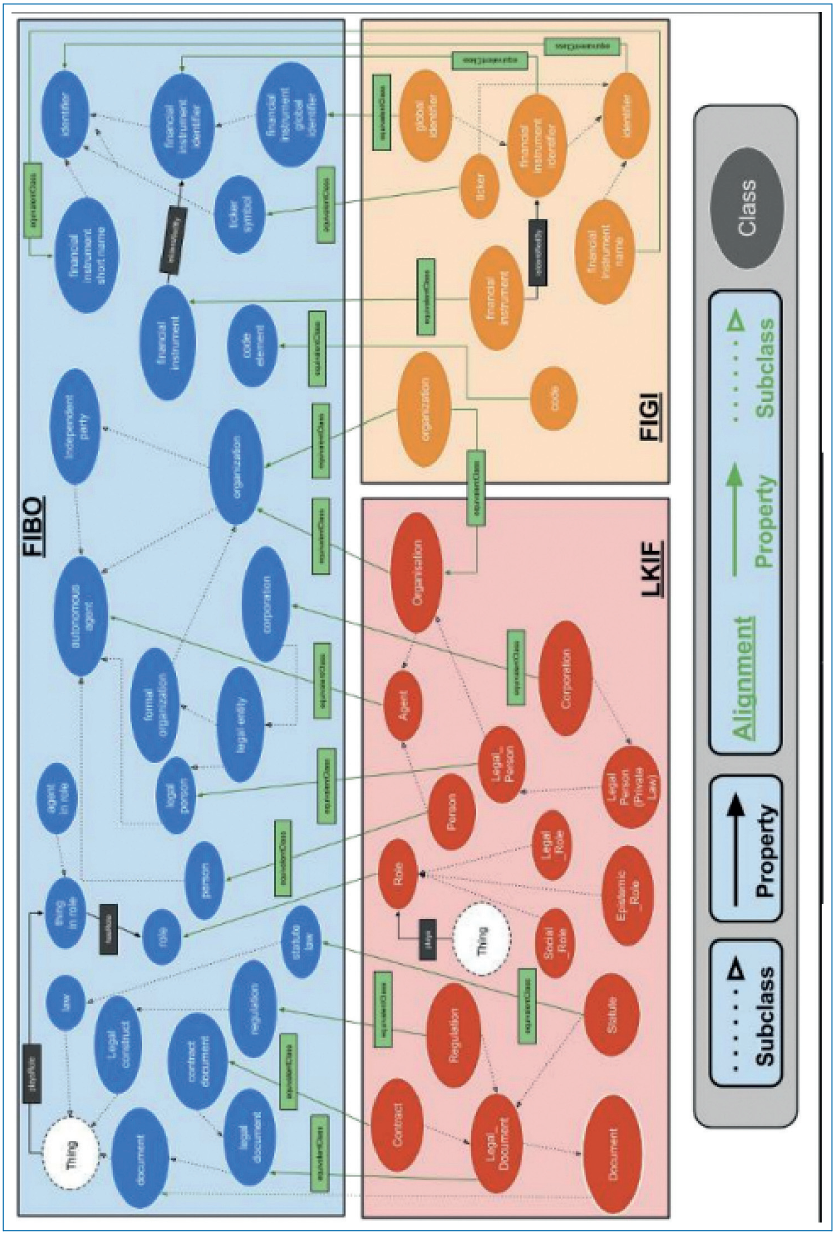

FIBO
FIGI
LKIF
identifier
financial instrument identifier
financial instrument global identifier
ticker symbol
financial instrument short name
financial instrument
code element
independent party
organization
autonomous agent
corporation
formal organization
legal entity
legal person
agent in role
person
thing in role
role
law
statute law
regulation
Legal construct
contract document
legal document
document
Thing
global identifier
ticker
identifier
financial instrument name
code
Organization
Agent
Corporation
Legal_Person
Person
Legal Person (Private Law)
Legal_Role
Role
Epistemic_Role
Social_Role
Statute
Regulation
Contract
Legal_Document
Document
plays
equivalentClass
isIdentifiedBy
Subclass
Property
Alignment
Class

Figure 4.4 INFINITECH core.

The INFINITECH core model and data pack define a lingua franca necessary to minimize the shortcomings of fragmented data from distinct data silos while harmonizing the data organization and knowledge representation within enterprises. In particular, the financial sector is covered by using FIBO as reference model while insurance sector is covered by LKIF. Furthermore, data pack ontologies and models for Internet- of-Things (IoT) derived from FIESTA-IoT or OpenIoT, are also considered, in order to consider one of the technologies that is driving the digital transformation where data are provided by ubiquitous devices. FIESTA-IoT provides tools, techniques, processes and best practices enabling IoT testbed/platforms operators to interconnect their facilities in an interoperable way based upon cutting edge semantics-based solutions. [INFINTECH Technical Report D1.6]

4.1.5 Example-TAHO (Traffic Analysis Hub Ontology)

TAH have developed a unique platform called the Traffic Analysis Hub (TA Hub), which ingests vast amounts of media content from diverse sources and builds an interactive map of people trafficking routes and hotspots. This platform is currently being used by many NGOs, Enforcement Agencies, and Financial Institutions to identify where they can best focus efforts in identifying and exposing the criminal gangs behind the trafficking. The TA Hub prototype runs in a secure IBM Cloud environment that was designed to meet the security needs of these partners, and it includes IBM's Watson AI and other analytical tools that analyze blended data to uncover Trafficking hot spots and routes that have not been evident before. The TAH tools are also pulling in and making sense of open source data, including thousands of daily public news feeds, to augment the data contributed by consortium partners, and to develop predictive capabilities in the future. [D6.3]

Pilot #3 leverages innovative INFINITECH technologies/ components in order to meet its objectives. The pilot participants collaborate to develop an AI driven capability using KYC/KYB methods and semantic technologies over transactional data generated by the financial activities that identifies money-related profiles based on the data generated. Data profiles then can be associated with human profiles based on their financial activity. These profiles will be built into the Watson AI engine and will be combined with existing technology and data sourced from the TAH human trafficking platform. The results will produce a complete picture of people's profile, people trafficking routes and the corresponding money flows back to the criminal organizations. [D6.3]

Pilot #3 implements KYC/KYB methodologies and uses emerging technologies investigating business innovation opportunities, as well as technology innovation for the banking sector and exploring the ways to reduce constraints that limit the development of new sharing data services. The Pilot participants are primarily financial institutions and organizations, investigating the introduction of data sharing capability to facilitate improvement of core banking and business capability alike the improvement of financial services. [D6.3]

The pilot proposes that KYC/KYB, a service that is well known and extensively used in financial and insurance companies when onboarding new customers, can benefit from the ability to share data securely and effectively between inter-bank departments, banks and also external services. The Pilot looks after identifying data patterns that can be related to unlawful activities and innovate by looking at the potential for fighting against human trafficking activities. KYC/KYB data sharing is a key characteristic for improving the financial sector, particularly now with the advent of FinTech companies where more people are investing using online platforms making more disruptive businesses, banking sectors need to enable the creation of new products and services. [D6.3]

Pilot #3 target is to build data profiles then can be associated, by use of INFINITECH semantic technologies, to human profiles based on their financial activity. These profiles named Red Flag Typologies will be built into an AI engine and will be combined with existing technology and data sourced from the Traffik Analysus Hub (TAH) human trafficking platform. The expected results are a capability to produce a complete financial profile of people that may incur on illegal activities, trafficking routes and the corresponding money flows back to the criminal organizations. The adoption of the INFINITECH Way Foundations by Pilot #3 application offers numerous business innovations. [D6.3]

Current Overall Application Scenario(s) for business innovation focuses on: [D6.3]

1. Money flows detection based on transactional data, this is always an interesting topic in the area of finance and banking transactions.
2. The identification of abnormal operations, which in today's banking systems is a trivial activity but the identification and detection of current money data flows and its traceability across different banking entities and financial institutions and organizations is yet a challenge and particularly when those transactions are associated to unlawful operations.
3. Detection of potential human trafficking activities, human trafficking is one of the fastest growing crimes in the world today, representing a $150 billion industry, and infiltrating supply chains at many levels.
4. Better understanding in how to fight and disrupt human trafficking crimes and put marked and indicators on transactions and/or operations that can be identified and thus end to the misery suffered by its victims.

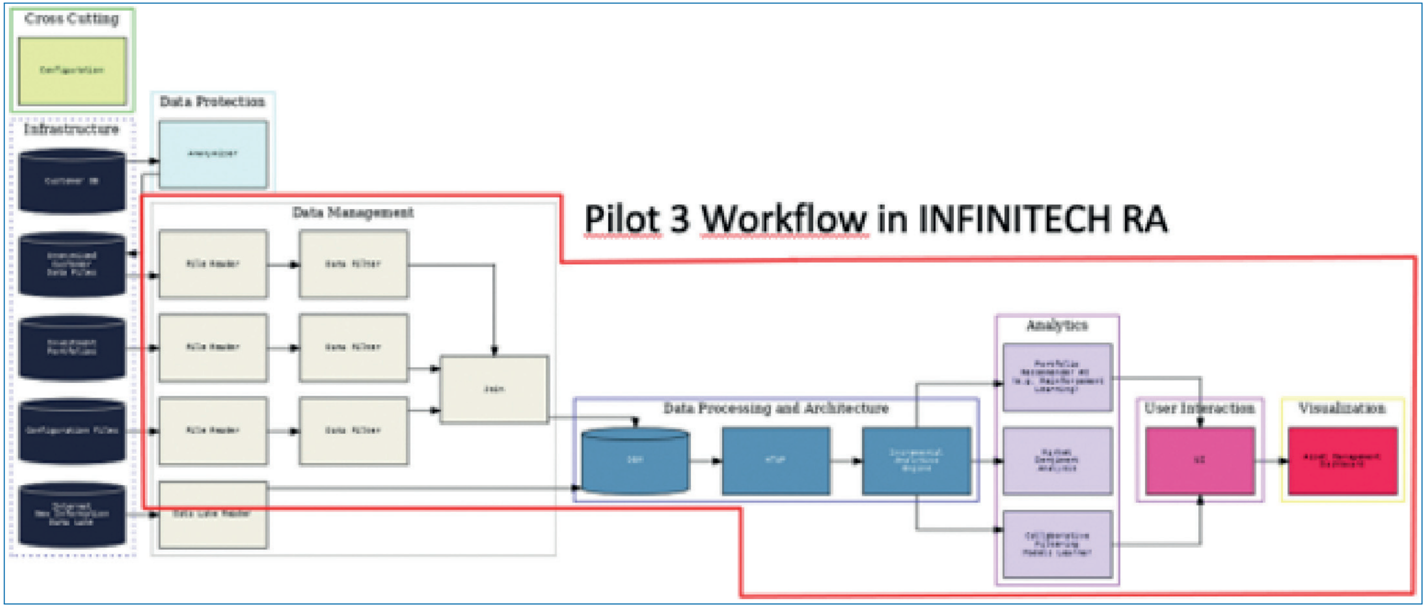

Cross Cutting
Infrastructure
Data Protection
Data Management
File Reader
Data Filter
File Reader
Data Filter
File Reader
Data Filter
Join
Data Lake Reader
Pilot 3 Workflow in INFINITECH RA
Data Processing and Architecture
Analytics
User Interaction
UI
Visualization

Figure 4.5 Pilot #3 designed architecture according to INFINITECH RA.

Figure 4.5 illustrates pilot #3 designed architecture according to INFINITECH-RA. Similar to the workflow defined in INFINITECH-RA section, here we have infrastructures, data protection, data management, data processing and architecture, data analytics, user interface, and visualization. [D2.13 to D2.15]

This innovation defines the Terminology used in INFINITECH Ecosystem. Vocabularies for the INFINITECH Ecosystem. Establish the Basis for the different taxonomies. i.e. TAHO Inclusion in the INFINITECH White Paper on The INFINITECH Readiness Level (IRL).[WP4 integrated presentation assets]

Description:

- This innovation defines the Terminology used in INFINITECH Ecosystem.
- Vocabularies for the INFINITECH Ecosystem.
- Establish the Basis for the different taxonomies. i.e. TAHO.
- Inclusion in the INFINITECH White Paper on IRL.

Business Value:

- Continuous availability of operations.
- Interoperability and Data Exchange in financial and insurance sectors.

Target Market:

- Banking, Financial Services, Insurance, FinTechs, Insurance Techs.

Ownership:

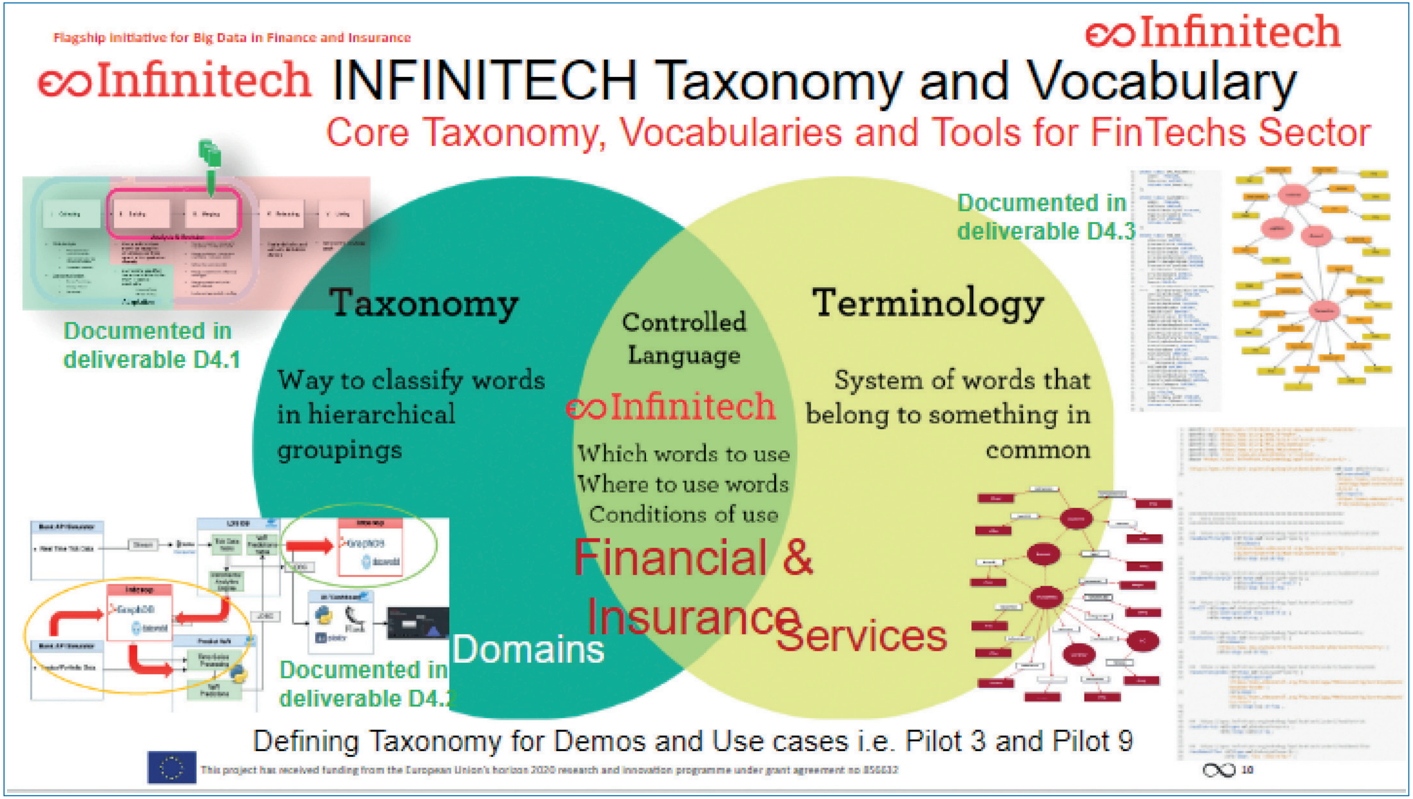
- NUIG and INOVA released under Creative Commons.

Figure 4.6 shows INFINITECH taxonomy and vocabulary and tools for fintech sector. It shows taxonomy which is a way to classify words in hierarchical grouping, controlled language which defines which words to use where to use the word and its condition of use, terminology which is a system of words that belongs to something in common.

4.1.6 Terminology Used in INFINITECH

Definitions used in INFINITECH Ecosystem and Contributions to the INFINITECH Readiness Level (IRL).

- Asset – A technology implemented or extended in the context of INFINITECH project.



Flagship Initiative for Big Data in Finance and Insurance

Infinitech

Infinitech INFINITECH Taxonomy and Vocabulary

Core Taxonomy, Vocabularies and Tools for FinTechs Sector

Documented in deliverable D4.1

Taxonomy

Way to classify words in hierarchical groupings

Controlled Language

Infinitech

Which words to use
Where to use words
Conditions of use

Terminology

System of words that belong to something in common

Documented in deliverable D4.3

Financial & Insurance Services

Domains

Documented in deliverable D4.2

Defining Taxonomy for Demos and Use cases i.e. Pilot 3 and Pilot 9

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement no 856632

Figure 4.6 INFINITECH taxonomy and vocabulary.

- PoC – An Example (Slides, Diagram, Video, Demo) How an INFINITECH Technology works.
- DEMOS – An implemented Example in how a Pilot potentially could use the PoC.
- Case Studies – The concept(s) or use cases from a Pilot behind a DEMO.
- Exploitation – The potential of a Case Study being implemented as part of a INFINITECH PoC.
- Integration – The DEMO showing/using INFINITECH technology/technologies.
- Deployment – A DEMO installed/instantiated in an INFINITECH Server or at Stakeholder Infrastructure.
- Operation – A DEMO running in INFINITECH Server or Stakeholder Infrastructure.
- Marketplace – INFINITECH Server or Stakeholder Infrastructure showing assets.
- Pilot – All the above + End Users.

DOI: 10.1561/9781638282297.ch5

Chapter 5

INFINITECH Technologies, Data and Processes

5.1 INFINITECH Technologies, Data, and Processes

This section summarizes technology specifications of the building blocks that will be used within the pilots and in particular in the pilots' sandboxes of the INFINITECH Project. It is conceived as a reference resource of information for the entire project about the components used and/or developed within the project. We detail the BigData/IoT technological building blocks that will be developed in the scope of the project, in the areas of data management, semantic interoperability, cost-effective real-time BigData analytics, elastic cloud storage, integrated (declarative) data querying, AI/ML algorithms and more.

This section presents the available tools and applications owned by the Consortium's Partners that constitute the background of the technologies exploited in the Project. These BACKGROUND Technologies can be the basis to build up other components for INFINITECH, by improving them and increasing the related TRL. The technologies are identified after a process/reasoning based on the input provided by the Consortium Partners. For each tool or application, the following information are provided: [D2.5–D2.6]

- Title/Name as a short description of tool/platform.
- Description in particular the characteristics and/or technology the component is based on.
- Documentation or detailed references to links/demo environment.
- ACRONYM of the Company or Partner in the project who has developed tool/platform and is owner.
- TRL Technology Readiness Level current and expected level at the end of the INFINITECH project.
- Ideas for enhancements to be done in the course of the project.
- License schema (e.g. Proprietary, GPL, Apache, MIT, …).
- Pilots (usually referred as P01, P02, etc).

5.2 Technologies

The following is a list of the possible technologies that may be useful in the fintech and insurance tech sectors, most of them are referred as INFINIETCH technologies as they are included in some of the processes in the pilots while others are included by its relevance on the sector. This list does not aim to be extensive but includes basic descriptions for a general public understanding about all the relevant technologies as a summary.

5.2.1 Data Ingestion

This technology is tasked with the capture, homogenization, distribution and storage of the datasets that support the connected car pilot. It involves the design and implementation of the IoT agents that adapt the data available from the Vehicle, as well as deploying the necessary modules for data storage and distribution (FIWARE Orion, FIWARE Cygnus, etc). As the data sources size increases, the technology will have to become more robust by polishing the performance of the deployed IoT agents.

5.2.2 Data Protection Orchestrator (DPO)

It is an enabler for embedding and automating the assurance of security and privacy by design and by default in heterogeneous and complex business flow. It orchestrates various privacy and security management functions (such as access control, encryption and anonymization). It will be used as part of the data governance Framework of the Project, and towards establishing the regulatory compliance tools in the project's sandboxes. It requires Swagger specification of the components (PETs) that will be called via REST by DPO There will be developed the business flows to address the specific communication with the components.

5.2.3 Digital User Onboarding System (DUOS)

It is a solution for dealing with virtual identities in a mobile device. It provides remote user registration using eID or passport. It is needed to use eIDs issued by European National authorities according to the EU eID schemas: eID cards and Passports There will be implemented the needed improvements regarding integration with end users application (Bank application) that needs user authentication.

5.2.4 EASIER-AI

EASIER-AI is an Hybrid (Cloud/Edge) platform that facilitates to develop, measure, monitor and deploy your AI models. The platform facilitates the data science tasks and it is focused on working on Hybrid Infrastructure and exploiting data generated by IoT. The platform synchronizes Cloud and Edge, keeping the Edge always up to date to run always the most accurate model. By including this tool in the INFINITECH project, we aim to feed it with new dataset sources, resulting in the development of new ML models for the platform.

5.2.5 Driver Profile Classifier

This technology is aimed at the use of high-quality vehicle data allows insurance companies to offer customized products. The application of supervised machine learning techniques is proposed to classify drivers' profiles which generates a customized insurance premium. The resulting model is then deployed with TensorFlow serving and integrated as part of the cloud platform with a wrapper, achieving an accuracy of 85.7%. The inclusion of this model in the project could bring an improvement on the data analysis of the datasets, bringing an improvement to the model accuracy.

5.2.6 Distributed Near-real-time HPC Processing and Exchange of IoT Streaming Data

AI algorithms' optimization exploiting GPUs capabilities. Usage of GPU processors to enhance AI algorithms performance and reinforce CPU capabilities.

5.2.7 Botakis Chatbot Development Network

A tool for rapid development of chatbots applications, which will be used for the development of chatbots features in the INFINITECH pilots (i.e. notably the LIB, BOC and NBG led pilots). Enhancements expecting to be achieved for Botakis Chatbot Platform, based on INFINITECH pilots (i.e. notably the GFT and NBG led pilots) : – Built-in dialogs that utilize and be integrated with existing NLP frameworks (open or proprietary) provided by partners or every interested party – Powerful dialog system with dialogs that are isolated and composable. – Built-in prompts for simple things like Yes/No, strings, numbers, enumerations.

5.2.8 Crowdpolicy Open (Innovation) Banking Solution

Crowdpolicy Open (innovation) banking platform is a set of predefined and customizable banking web services and data models integrated with our own API Manager that supports access control, monitoring and authentication. Our solution puts the bank (or any monetary financial institution) in control of the third-party partner relation. The solution is full PSD2 & GDPR Compliance. Enhancement aim through INFINITECH project are : – technology scale-up is to democratize the use and exploitation of open banking APIs even for users with no development skills, building fintech software development kits. – implement a complete programmable framework to integrate different services and apis using protocols by providing similar user experience as zapier, “yahoo pipes” and “IFTTT”. The main objective at the innovation perspective is to provide a graphical user interface for building data and fintech services mashups that aggregate open banking APIs, open available data sets and rules and creating Web based apps from various sources, and publishing those apps.

5.2.9 AI-Engine-for-Psychometric-Profiling and Personalization

AI-driven engine to extract four categories of behavioral features, grouped according to the type of spending behavior they capture: (i) overall spending behavior, (ii) temporal spending behavior, (iii) category-related spending behavior, and (iv) customer category profile. The engine uses these features to predict the personality of customers.

5.2.10 Open Source AI/ML Frameworks

These frameworks facilitate the development of AI/ML based tools, which shall be applied to Financial Crime and Fraud, e.g. on so called Instant Loans. Today a number of open source tools for AI/ML development are available. The AI/ML community is progressing these technologies dynamically. This way it provides the basis for solution development and facilitate the specific solution of a wide range of business problems as in INFINITECH. This way, these open source tools provide the foundation for development towards off-the-shelf modules being part of the INFINITECH-RA.

5.2.11 Data Layer – REST API

A Data Layer to support Security Data Model with REST API based on a not relational database (MongoDB). Supports heterogeneous sources. Developed upon FLASK-Python3 framework and dockerized to be deployed on Kubernetes infrastructure. This tool aims to complete the wrap with standard I/O.

5.2.12 Terrier Information Retrieval Platform

Search Engine for BigData sets that offers integration with Spark for distributed processing. We plan to extend Terrier with a new open source module that combines real-time data stream ingestion (via Apache Flink) with distributed database access (via LeanXcale) for real-time data indexing and updating from multiple sources, within WP3 and WP5. We will also expand Terrier with enhanced Python integration, allowing easier use from common data science pipelines, such as those involving Pandas. This technology can be used for tasks such as searching/sampling financial product portfolios, user profiles or for providing recommendations.

5.2.13 Anonymization Tool

A tool that anonymizes data in order to preserve privacy. It also provides metrics that allow to measure the risk of the anonymized data and the impact of the anonymization process on the utility of the data. The tool will be used in pilot #11 and pilot #12. The component needs a specific configuration/development for each pilot in which it is used.

5.2.14 Polyglot Database Management System

The LXS DBMS is a polystore database that provides access to different and heterogeneous datastores via a common interface. It allows for the data user to submit a query, whose scan operators can request data that are stored in external datastores and combine their intermediate results with data coming from other sources, either LXS internal datastore or others. For instance, a JOIN operator might require to JOIN table A (resigned in LXS) and table B (resigned in a MongoDB or a Hadoop

DataLake). At this phase, there is a support for a limited target datastores, for Proof-of-Concept of the prototype. Moreover, the user has to write queries for the target datastores in the specific dialect. What is more, JOIN operations are not efficient, as they require all data resigned in an external datastore to be retrieved in the query engine level. The enhancements that are planned to be implemented are the following: (1) Provide support for a variety of different datastores, according to the needs of the pilots. (2) introduce a novel SQL-like query language, so that the data user can write a query in a seamless way, and let the polystore interprete it to the target datastore. (3) improve the query engine in order to take into account scan operations (mostly part of the one of the JOIN arguments) that need to retrieve date form an external datastore. This will require the query optimizer to explore equivalent operation graphs based on the nature of the target polystore, the operator to be able to push down operations to the target store, and the query processor to set up the corresponding data pipelines during the execution of the query plan. It is not clear at this moment, if the corresponding wrappers/connectors for each target datastore will be able to retrieve statistics to feed the query optimizer. This depends on the target datastores that will be used by the pilots, and their capabilities to expose this type of information.

5.2.15 HTAP Database for the Financial and Insurance Sector

An ultra-scalable SQL Database and real-time big data platform that revolutionize the business database management systems by introducing the next generation business database that can scale in any of the three Vs of Big Data (Volume, Velocity and Variety). In more technical details, it provides an ultra-scalable transactional management system that can scale out to 100s of nodes, which is typically a bottleneck in traditional database systems that provide transactional semantics, while on the same time, is full SQL compatible and ensures all ACID transactional properties. It additionally exposes an interface for direct access of its key-value storage engine, thus providing a dual access without downgrade transactional semantics. It offers OLTP and OLAP integration, thus providing support for HTAP that allows for analytical queries over operational data, which realizes the concept of real time business intelligence. Finally, it enables for the execution of polyglot query processing across different and heterogeneous data sources. Modules of the database will be used for implemented all building blocks of INFINITECH that are related to data management. LXS background technology will be enhanced in order to support the data management building blocks: Mainly, it will be enhanced in order to comply with the requirements for HTAP support, and to be compliant with the target data sources that need to be accessed via the polyglot mechanism. Moreover, it will provide support for real time query processing, enabling queries that combine both streaming data with data at rest. Finally, it will provide the support for incremental and parallel query analytics. However, at this phase of the project, it is not yet defined the exact technical details regarding these enhancements.

5.2.16 Natural Language Processing for Real-time, High-accuracy Credit Risk Assessment

ReportBrain's NLP functionalities: – Provide real-time structured feeds on risk-assessment worthy information sourced from the news – Interlink entities, updates and maintains knowledge graphs in real-time with all the interlinking of entities – Use "visual" algorithms to collect and analyze the news in real-time in 65 languages 24/7 – Classify articles in real-time by their content (politics, business etc.) – Use its own models, identifies entities (organizations, persons and locations) in real-time. By using the AI enhancements that will be developed by Reportbrain in financial services & insurance sectors, users will be able to add an extra, yet orthocanonical feed to their existing credit rating models that will provide a real-time understanding of the world and more specifically on what's if happening with specific entities (organizations & persons) of interest.

5.2.17 Machine Learning Algorithms for Health Related Data

SILO has implemented such approaches in different health-related projects such as CrowdHEALT. SILO to make suggestions for the enhancement of the Platform.

5.2.18 Wenalyze Big Data Analytics Platform

Platform that collect and process information from multiple open data sources regarding SMEs and apply cognitive algorithm to detect risk and changes in financial needs. The tool will be use in pilot 13.

5.2.19 Octopush Geospatial Enabling Framework

Octopush is a geospatial enabling framework, developed by AgroApps, allowing the collection, pre-processing, post-processing and distribution of geospatial data products and services, either referred to remote sensing (satellite, drones) acquisitions or multidimensional data outputs from numerical simulations. Octopush allows users to have access through a centralized access point to decentralized services, while Octopush SDK enables IT developers, to easily adapt or expand the provided geospatial services. Octopush was created by AgroApps aiming to adders the company/operations and services need for a modular system, independent from any third-party service provider (excluding those offering raw data like Copernicus, NASA etc.), a framework that will be easily adapted to the market needs and follow the service-oriented business model of the company. Octopush is the baseline framework that addresses the AgI companies need in geospatial information, either through the development of new services and data models or the adaptation of the existing ones. Some of the services that are currently offered by Octopush are: a significant number of Vegetation Indices derived from Optical and SAR imagery and crop specific biophysical parameters including leaf area index, chlorophyll content and above ground biomass; crop specific yield estimation; farm management information services such as irrigation scheduling and variable rate fertilization; weather-driven models of possible pests and diseases outbreaks; high resolution weather forecasts and specific agrometeorological parameters; crop damage assessment services.

5.2.20 AgroApps Weather Intelligence Engine

Weather Intelligence Engine, developed by AgroApps, is a numerical weather prediction and atmospheric data assimilation processing chain, based on the WRF numerical weather prediction model. Weather Intelligence Engine is producing operationally all the needed weather data products (Near-real-time, medium-range weather forecasting, subseasonal to seasonal forecasts) by AgroApps offered services. Weather Intelligence Service could take advantage of the available INFINITECH HPC resources, and pilot test a hybrid ensemble data assimilation scheme in convective scales.

5.2.21 Sentiment Analysis Tool

Reportbrain Sentiment analysis tool uses application programming interface (API) calls to search existing news article index – Elastic Search Index. The results of the search are processed in Reportbrain's Sentiment Analyzer and the outcome is returned to the caller as a REST API response. This means that articles that are requested by an authorized caller are evaluated in real-time for sentiment and returned to the API caller. Sentiment evaluation describes sentiment as 0 for neutral, -1 for negative or $+1$ for positive. The purpose of the RB News Article Sentiment API is to provide the sentiment of articles selected by the user via a Query. In particular, the RB News Article Sentiment API provides several fields, that could be used as filters to query on RB platform. These are: content, language and date. The user can retrieve the sentiment analysis to personalize portfolios of their clients, using valuable insights of positive/negative/neutral evaluation of news articles about the entity of interest.

5.2.22 Partitioned and Distributed Transaction Graphs

Ethereum and bitcoin public blockchain transaction datasets. BOUN to make suggestions for the enhancement of the Platform.

5.2.23 ALIDA: A Micro-service Based Platform for Composition, Deployment and Execution of BDA Applications

A micro-service based platform for composition, deployment, optimization, execution and monitoring of big data analytics workflows (covering ingestion, preparation, analysis and visualization). It is designed and developed on top of the most cutting-edge open source Big Data technologies and framework.

5.2.24 Text Analysis Tool

Reportbrain text analysis tool generates insights from both structured, semi-structured and unstructured text data using natural language processing (NLP). Such insights include sentiment analysis, key phrases, language, and entities, among others. The Reportbrain Analytics Engine uses advanced parallel processing and combines complex NLP tasks in real- time to produce desired results. Text analysis tool allows companies to better understand all types of data they are interested in. After deep analysis of data lake performed by Text Analysis tool, user gain valuable insights about entity that was used in a query. Performing the same process manually would require tremendous amount of effort and time. The final objective for using the tool is that knowledge is provided, and not-relevant information are ignored.

5.2.25 Blockchain Tokenization

Hyperledger Fabric blockchain support for tokens. Enhance Hyperledger Fabric with tokenization capabilities for digital trading of assets.

5.2.26 Healthentia LifeSciences BigData Platform

BigData platform providing data sources aggregation and management, as well as tools for analytics and visualization. It will be used in the IoT-based Life Insurance pilot (ARI/RRD Lead pilot). Re-purpose the platform to support the insourance-related monitoring of people for pilot 12.

5.2.27 Event-registry

Event Registry (ER) is real-time cross-lingual global media monitoring service for modeling global social dynamics (eventregistry.org) developed by the JSI. ER aggregates and analyses news content for over 120,000 news sources published globally in 100+ languages. Events mentioned in the news are identified and relevant information about them is automatically extracted and stored in a searchable form. The data can be accessed directly on the platform or via the API. ER supports various analytics including deep analytics of the events and correlations between events, extracted entities and financial data extracted from the main financial indexes. Eventregistry could be enriched with new insights of the potential use scenarios and would benefit from them to expand the current offering to the fintech industry. In addition, better insight will be given to the team in order to develop further analytics.

5.2.28 Stream Story

Stream-Story multi-resolution modelling and explanation of (possibly real-time) streaming data: (1) Exploratory data mining – A system for the analysis of multivariate time series. It computes and visualizes a hierarchical Markov chain model which captures the qualitative behavior of the systems' dynamics.; (2) Multi-scale representation – The hierarchical model allows users to interactively find suitable scales for interpreting the data; (3) Real-time monitoring Visualizes streaming data by mapping it to the hierarchical model. It can provide predictions and alarms for different behavior.

5.2.29 Qminer

QMiner is an analytics platform for large-scale real-time streams containing structured and unstructured data. It is designed for scaling to millions of data points on high-end commodity hardware, providing efficient storage, retrieval and analytics mechanisms with real-time response.

5.2.30 SSC – Super Stream Collider, a Multiformat Data Management and Query System

The SSC enables distributed cloud- based high-performance processing of semantically linked streams i.e. it is an enabler for semantic analytics. It will be used for analytics over semantically unified/interoperable streams (in WP5), as well as in the KYC and customer- centric services pilot (BOI-led pilot).

5.2.31 The Global Engine with Neural Network Intelligence (GENNI)

An AI-Powered Neural network-powered Engine executing DL (deep neural networks) algorithms over semantically annotated streams. It can be used in pilots for customer centric analytics.

5.2.32 Data Check-in Mechanism

A sophisticated data check-in mechanism that is enabling the preparation and uploading of the data provider's (public or confidential) datasets in the cloud platform that is one of the results of the ICARUS H2020 project. The data check-in mechanism is deployed on the premises of the data provider as a stand-alone desktop application and receives as input a list of data check-in jobs that incorporate a set of instructions with all the actions that will be performed on a specific dataset, residing on the local storage of running operating system, in order to enable the data preparation and uploading of new datasets in a secure manner. Internally, the mechanism handles the orchestration and execution of the designed instructions with the use of incorporated (micro) services for the: (a) data mapping of data source entities to the designed common data schema, (b) data cleaning operations on the data source entities, (c) the anonymization operations on the data source entities and (d) the encryption of the data source entities . This list of (micro) services is expandable based on the needs of each platform. The data check-in mechanism is offered in the form of a local client for all OS (Mac, Linux, Windows) and is designed and developed using the latest technologies for desktop apps with the aim to offer end-to-end security on the data preparation and data upload tasks. The specific technology served as the basis for the design and implementation of the INFINITECH Data Collection component.

5.2.33 IoT-Catalogue

The 'IoT Catalogue' is a web-based catalogue from where to pick & choose IoT solutions; it is an explorer of innovations in IoT applications and technologies. It aims to be single entry point of support to IoT developers/integrators/advisors/end-users in the process of identifying and selecting IoT technologies (ranging from complete end-to-end solutions to tools and components/parts) but also inspecting a wide set of IoT use-cases, their validations, associated contact persons/organizations, detailed characterization (value propositions, ICT problems, functions, target, domain), supporting technological solutions, and much more.

5.2.34 Analytics Library

In the scope of ATMOSPHERE (Adaptive, Trustworthy, Manageable, Orchestrated, Secure Privacy-assuring Hybrid, Ecosystem for REsilient Cloud Computing) project, the UPRC team, focused on the delivery of the library of services, which can be utilized as a baseline for the INFINITECH library. (WP5). Update the library to include metadata relevant to security and privacy constraints of the INFINITECH algorithms to be made available through the library.

5.2.35 Catalogue of Objects

During the 5GTANGO and MATILDA EU Projects, both aiming to enable the flexible programmability of 5G networks and to devise and realize a radical shift in the development of software for 5G-ready applications, UPRC team contributed to the market platform and to the catalogue of services & functions, respectively. Therefore, the relative outcomes and the solutions developed by UPRC for these projects will be utilized as a baseline for the INFINITECH's marketplace, integrating with UNP's IoT-catalogue. Extend to support a variety of assets (e.g. datasets, models, etc) as well as to support composite assets (i.e. analytics pipelines).

5.2.36 Portfolio Optimization Tool

Privé goal is to automatically construct an optimized portfolio by modelizing the investment advisory and decision process using a certain level of AI procedures. The result is a tailored portfolio for each individual investor. All functionalities shall be also available via API access, which should be a kind of "FinTech-as-a-service" (FaaS). Our target market are all financial services intermediaries who provide advisory and wealth management services. Hence Banks, Insurers, Insurance Brokers, EAMs, Securities and Brokerage firms are the target customers for this solution.

5.2.37 Sentiment Analysis for Financial News

This is a REST API taking as input financial news in text form and performs Sentiment Analysis on the given text. Additionally, the utilised model is be retrained periodically retrieving historical news from LXS database. It enables classification of financial news according to their impact (i.e. positive, neutral, negative) on a given portfolio. This tool can be used in parallel with other quantitative metrics in order to provide a comprehensive risk assessment in the stakeholders (traders, risk managers etc.).

5.2.38 AI Model for VaR Prediction

This component is a REST API which predicts Value-at-Risk and Expected Shortfall of several financial Portfolios, utilising both well-established and innovative techniques. AI model for VaR prediction takes as input both the asset prices and the current trading position and derives the VaR and ES estimates at 95% and 99% confidence level utilizing three different models. The estimation procedures are repeated every minute to take into account the most recently available data providing risk assessment in (near) real time. Our target market are the institutions which are exposed to market risk such as commercial and investment banks, insurance companies and institutional investors.

5.2.39 INFINITECH Components Specifications

The INFINITECH Reference Architecture (IRA) [Deliverables D2.13, D2.14], defines a set of layers and allows flexible workflows out of data processing modules, called components that performs specific data transformations. A workflow usually consumes one set of data (sources) to produce another set of data (destinations). A component can be identified by the Input/Output interfaces and the functionality (transformation) provided at the edge. In that respect, a data component can be considered as a black box that can be replaced by any technological implementation that performs at the edge the same transformation.

Figure 5.1 depicts the different reference layers of the IRA along with examples of data components. Scope of the section is to identify existing and desired data components that will support the design, the development and deployment of the Pilots' sandboxes.

An INFINITECH solution (sandbox) can be built organizing components in a workflow (sometimes referred to data pipeline) to accomplish a complex transformation from one set of data sources to another set of data solving the business case. Components should therefore be interoperable, with clear interfaces and perform a clear function over the data. The rest of this section finally describes the existing and to be developed components of the project. Further progress in the project will provide the needed details for implementation and deployment.

5.2.40 Data Component Description

In the following a first list of the identified components, are described using a standard template described in the table below.

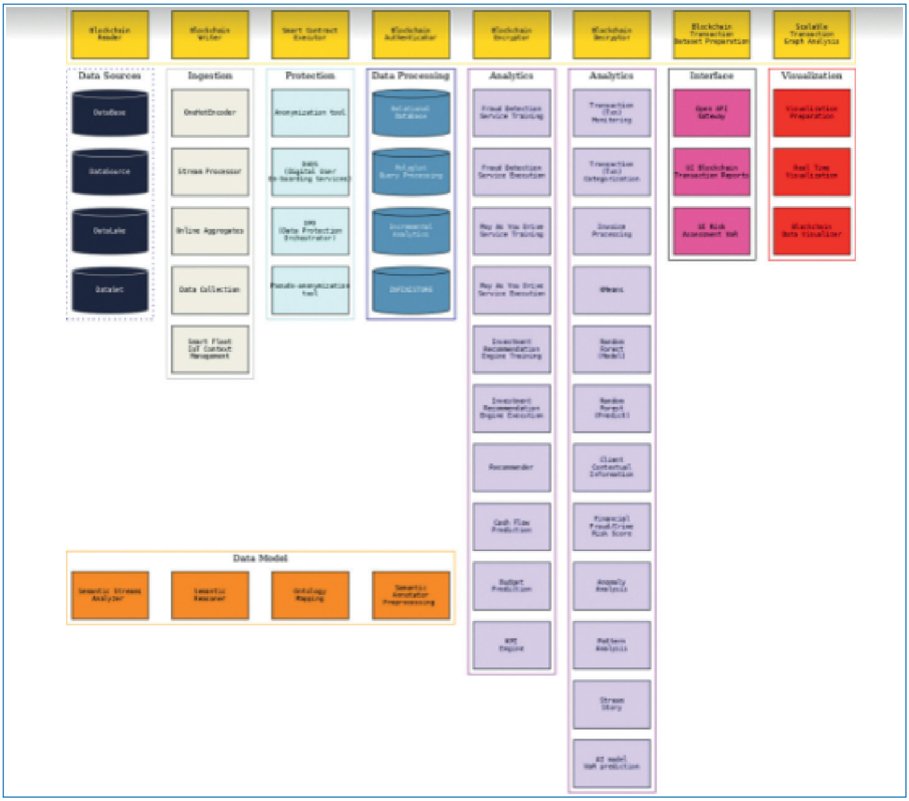

Blockchain Reader
Blockchain Writer
Smart Contract Executor
Blockchain Authenticator
Blockchain Encryptor
Blockchain Decryptor
Blockchain Transaction Dataset Preparation
Scalable Transaction Graph Analysis
Data Sources
DataBase
DataSource
DataLake
DataSet
Ingestion
OneHotEncoder
Stream Processor
Online Aggregates
Data Collection
Smart Fleet IoT Context Management
Protection
Anonymization tool
DUOS (Digital User On-boarding Services)
DPO (Data Protection Orchestrator)
Pseudo-anonymization tool
Data Processing
Relational Database
Polyglot Query Processing
Incremental Analytics
INFINISTORE
Analytics
Fraud Detection Service Training
Fraud Detection Service Execution
Pay As You Drive Service Training
Pay As You Drive Service Execution
Investment Recommendation Engine Training
Investment Recommendation Engine Execution
Recommender
Cash Flow Prediction
Budget Prediction
KPI Engine
Analytics
Transaction (Tx) Monitoring
Transaction (Tx) Categorization
Invoice Processing
KMeans
Random Forest (Model)
Random Forest (Predict)
Client Contextual Information
Financial Fraud/Crime Risk Score
Anomaly Analysis
Pattern Analysis
Stream Story
AI model VaR prediction
Interface
Open API Gateway
UI Blockchain Transaction Reports
UI Risk Assessment VaR
Visualization
Visualization Preparation
Real Time Visualization
Blockchain Data Visualizer
Data Model
Semantic Streams Analyzer
Semantic Reasoner
Ontology Mapping
Semantic Annotator Preprocessing

Figure 5.1. Logical schema of data processing components in INFINITECH.

5.3 Data Components

5.3.1 Relational Database

This component consists in the central data repository of the platform. It enables transactional semantics and provides capabilities for query processing based on standard SQL statements. It can scale out on the runtime while continuing serving operational workloads. It can support analytical processing in combination with operational data modifications with the level of isolation to be snapshot isolation. That is, it enables real-time business analytics. The data repository is written in Java and C, and provides support for JDBC, ODBC and python drivers. It runs on K8S cluster.

5.3.2 Polyglot Query Processing

This component enables the query execution over more than one datastore in seamless manner. The data user can submit a single statement and let this component to execute the query by pushing it down to the target database. By doing this, it can process the data on premise, and retrieve only the results, thus it is convenient for cases where data cannot be loaded the platform and they need to be accessed from an external datastore. This component is implemented in Java. It is an extension of the central repository itself. In fact, it consists of a jar binary which is loaded in the classpath of the query engine of the data repository. As of that, it will be deployed as part of the query engine via K8S cluster.

5.3.3 Incremental Analytics

This component enables the query execution in an incremental fashion. The data user will be able to submit a continuous query to the datastore, which will be continuously and incrementally validated. This means that the initial results will be retrieved first, and as data arrives to the data repository, they will be validated against the submitted query, and if it validates the statement, it will be returned to the user. This component is implemented in Java. It is an extension of the central repository itself. In fact, it consists of a jar binary which is loaded in the class path of the query engine of the data repository. As of that, it will be deployed as part of the query engine via K8S cluster.

5.3.4 OneHotEncoder

Service to represent categorical variables as binary vectors. It is a BDA service registered in the ALIDA catalogue. It is a PySpark-based micro-service running on K8S Spark cluster mode, working as part of the ALIDA framework. In a nutshell, ALIDA is a Micro-service based platform for composition, deployment, optimization, execution and monitoring of pipelines of Big Data Analytics (BDA) services. ALIDA is a result of previous research activities developed by ENG. Currently, it is a work in progress. ALIDA offers a catalogue of BDA services (ingestion, preparation, analysis, visualization): user designs his own (stream/batch) pipeline by choosing the BDA services from it, indicates which Big Data set he wants to process, launches and monitors the execution of the pipeline and personalizes the results visualization by choosing from a set of available graphs, all this without worrying about having software developer skills or particular knowledge on big data technologies. This service is registered in ALIDA catalogue as Spring Boot Application containing the python code and its dependencies. After implementing the algorithm using Pyspark, creating the Dockerfile and pushing the new image inside a repository, this microservice is registered into the ALIDA catalogue through the GUI. Source: https://home.alidalab.it/.

5.3.5 Stream Processor

This component will provide streaming processing capabilities. The data user can declare continuous queries that will be executed over the data stream. It will also allow to combine streaming data with data at-rest and enable the storage of data streams even when injected in very high rates. This component will be based on Apache Flink and will be containerized to be deploy with a K8S cluster. This component can exploit the capabilities of the declarative real time analytics. This will be very useful in cases where the data user wants to calculate a value over a stream that will need to perform an expensive scan operation over a data table (i.e. compare the input with the overall average of a field in a table). As scan operations (and operations that will require a scan operation, as the average etc) have a complexity bigger then O(1), they are timely costly and cannot be executed in a stream. For that, the developer often caches that value and periodically updates the value. With the declarative real-time analytics, the data user is given the opportunity to declare such an analytical operation (i.e. the overall average) with an SQL fashion, and the query will be executed with a complexity of a get operation, which will allow these types of analytical operations to be included in a stream operation, providing to it the current average, with respect to data consistency and isolation in terms of ACID properties and transactional semantics.

5.3.6 Online Aggregates

This component allows for the execution of aggregate processing operators in an online manner. This way, the definition of the aggregate operations can be defined, and the result of the execution can be pre-calculated in an online manner, preserving data consistency and transactional semantics. When requested, the result of this execution can be retrieved with a GET operation, removing the need to scan the whole dataset. This component is based on the relation database component and extends its core storage and query engine. It runs on K8S cluster.

5.3.7 Data Collection

This component provides the data ingestion mechanism that: (a) enables the acquisition and retrieval of heterogeneous data from a variety of diverse data sources and data providers, (b) facilitates the data annotation of the retrieved data by enabling the mapping between the data entities included in the retrieved data and the provided by the data provider data model, (c) enables the design and execution of data cleaning operations towards the increase of the data quality of the retrieved data. This component is based on Java and Python programming languages. Additionally, Java Spring Boot, Flask, Pandas and NumPy frameworks and libraries are leveraged. This component is providing a highly configurable mechanism capable of addressing the various connectivity and communication challenges raised during data ingestion. Hence, the data provider is able to configure this mechanism in order to execute data collection pipelines that include data retrieval, data annotation and data cleaning operations which are tailored to their needs.

5.3.8 Anonymization Tool

The anonymization tool modifies data in order to preserve privacy. It is especially indicated in those cases where a dataset contains personal data and it has to be outsourced or shared with a third party. The tool includes different anonymization algorithms that aim at avoiding the appearances of data combinations that could lead to a possible re-identification of the data subjects. It also includes a set of privacy and utility metrics that allow to measure the risk that remains after anonymizing the dataset, and the impact of the anonymization process on the quality of the data. The tool is based on two modules that can be deployed as a Docker container. The tool requires to retrieve and store the raw and anonymized version of the data from a relational database.

5.3.9 DUOS (Digital User On-boarding Services)

Provides remote user registration using eID or passport, dealing with virtual identities in a mobile device. It uses various identity proofing and verification services that link new user eID creation (virtual or derived eID) with government issued e-ID. It verifies electronic data stored on chip and machine readable zone. Provides Flexible Multi-factor authentication for different users or identities. Different underlying licenses (Apache 2.0 license for MRZ reader and license to be decided for biometric checking).

5.3.10 DPO (Data Protection Orchestrator)

The DPO embeds and automats the assurance of security and privacy by design and by default in complex business flows. It provides orchestration of Privacy Enhancing Technologies and related services using BPM tools in order to integrate privacy or data protection perspectives into business processes. It can orchestrate any kind of REST service. The DPO interacts with the privacy expert that prepares the business flow in a BPMN file. The flow interacts with PETs such as anonymization.

5.3.11 Blockchain Reader

Fetches requested data from blockchain ledger. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavors of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer/Know Your Business, Asset Tokenization).

5.3.12 Blockchain Writer

Submits transactions on the blockchain ledger. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer/Know Your Business, Asset Tokenization).

5.3.13 Smart Contract Executor

Executes smart contracts on the blockchain ledger. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer/Know Your Business, Asset Tokenization).

5.3.14 Blockchain Data Visualizer

Queries and displays information about blocks, transactions, chaincodes and transaction families, network name, status and nodes list, organizations list and peers list.The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer/Know Your Business, Asset Tokenization).

5.3.15 Blockchain Authenticator

Grants access to specific channel(s) of the blockchain network.The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer/Know Your Business, Asset Tokenization).

5.3.16 Blockchain Encryptor

Encrypts the clients' sensitive data within the smart contract using AES256. The specific component is part of a Blockchain chaincode. As the chaincode is tightly connected to the business operation that is performed on top of different business objects, different flavours of the chaincode exist depending on the business use cases (Consent Management, Know Your Customer/Know Your Business, Asset Tokenization).

5.3.17 Blockchain Decryptor

Decrypts data retrieved from the blockchain ledger.

5.3.18 Blockchain Transaction Dataset Preparation Component

This component will be responsible for retrieving raw transaction blocks from the Bitcoin and Ethereum blockchains and parsing the blocks in order to extract Bitcoin, Ethereum and major token transactions. After retrieving all the blocks up until now, this component will be run periodically (e.g. once a week or as needed) to retrieve newly generated blockchain blocks during the period. Blockchain blocks must be retrieved from either nodes attached to the blockchain (e.g. running Parity) or from blockchain data supplier gateways (Google, Infura, Cloudflare). Web3.js Ethereum Javascript API and Consensys abi-decoder for smart contract call data parsing.

5.3.19 Scalable Transaction Graph Analysis Component

This component will be responsible for taking massive bitcoin and ethereum public transaction data. Since transaction graph size massive and growing, it will use parallel algorithms to achieve scalability. It will utilize graph and machine learning algorithms to analyse fraudulent transactions. Parallel processing (high performance computing) technologies. Message Passing Interface (MPI). Graph Partitioning Software such as Scotch and/or Metis. Distributed graph algorithms. Machine Learning.

5.3.20 Semantic Streams Analyzer

Semantic Streams Analyzer Middleware-Engine – SeSAME, The SeSAME component is a data mashup builder for the financial sector that can be used as a data processing component for your data management application, it enhances the capacity to process financial and insurance data in the form of batches and provides a single output, it is ideal when multiple sources have different data formats, it is built to be compatible with the most common data formats in the financial and Insurance sector i.e. FIBO, FIGI and LKIF and additionally it uses INFINITECH Core Graph Data Model to enhance performance. The SeSAME component is designed as a dataflow/workflow execution framework connecting various data input/outputs through the concept of pipelines for creating the data mashup. Conceptually, each financial operator has input data or streams and SeSAME provides one output data or stream. The multiple inputs can be used simultaneously while a single output in RDF is provided. Only the final operator of a workflow can return a format other than RDF, if necessary by defining and transforming the data into the desired format. The data Operators can be of three modes via APIs: API (1) a data acquisition operator is used to collect or receive data from data sources or gateways and can be pull-based or push-based. API (2) a stream processing operator defines stream processing functionalities in a declarative language, e.g., CQELS. API (3) a streaming operator streams the outputs of the final operator of a workflow to the consuming applications. In these three API modes operators of the data transformations and alignments can be done to produce a normalised RDF output format.

5.3.21 Semantic Reasoner

Enhanced Distributed Reasoner over FinTech Ontologies – EnDoRFIN Semantic Reasoner, The EnDoRFIN component is a tool for inferring knowledge from data streams, it uses some rules as conditions for defining logics conditions and as a result logical consequences are provided as outcomes. The inference rules are defined based on the most commonly used financial and insurance vocabularies i.e. FIBO, FIGI and LKIF and the way to process the rules is using APIs for defining the logical descriptions for the data applications it is introduced. This component allow the use of other languages but need to be upgraded to the target vocabulary and additionally the EnDoRFIN uses INFINITECH Core Graph Data Model to ensure the inference is applicable to all the involved domains in FINTECHs.

5.3.22 Ontology Mapping

INFINITECH Graph Data Model – Online Ontology Mapping Framework and Toolkit, The INFINITECH Graph Data Model is the set of online tools referring to the graphs, formats, vocabularies and ontologies used in the INFINITECH project. The INFINITECH Graph Data Model is provided in the form of a set of online accessible files, schemas and metadata model diagrams that represent the way the INFINITECH data can be organised and structured, it also contains the metadata in two different formats, json-ld and owl. The Ontology Mapping Ontologies section contains online machine-readable files both in OWL and JSON-LD format for online accessibility, both files are maintained and updated regularly to keep the latest version of the ontology files up to date using a versioning method.

5.3.23 Semantic Annotator-Preprocessing

Semantic Annotator-Middleware Preprocessing Layer for FinTechs – SAMPLe-FIN, The INFINITECH SAMPLe-Fin is the support online tool for transforming datasets into RDF- compatible format, beside the online available tool, a set of documentation is provided providing the necessary steps to transform data sets from any data-exchange format i.e. CSV, XLS, etc into RDF. This tool is provided as enabler for a semantic layer where enriched data can be processed more efficiently. INFINITECH SAMPLe-Fin is the mechanism INFINITECH uses for addressing cross domain and cross pilot Data Interoperability and Data Exchange and it also provides the pre-processing layer for the interoperability requirements in INFINITECH project.

5.3.24 Smart Fleet (IoT Context Management and Historical Data Component)

A FIWARE-Based framework designed to capture, homogenise, process and distribute real time traffic and smart vehicle's information (it will also allow other related context information). It will implement Pub/Sub mechanisms and support Geolocation and Time series tools. Additional tools to build custom dashboards will be included. Customised technological building blocks based on FIWARE architecture:Context Broker (Orion) [NGSI & ETSI NGSI-LD]: distribute context data between any possible combination of data-producers and data-consumersHistorical Data (Quantumleap) [ETSI NGSI-LD]: persistence of the data managed by the Context Broker.

5.3.25 Fraud Detection Service Training

This component trains the corresponding Machine Learning model that will assign a driver's profile and helps to identify the driver's behaviour using the provided driving routes and vehicle's technical data.

5.3.26 Fraud Detection Service Execution

This component provides the drivers' profile and classification (behaviour) for the last given driving route.

5.3.27 Pay As You Drive Service Training

This component trains the Machine Learning model that will classify the drivers' behavior while driving, according to the data collected from the driven cars (driving routes).

5.3.28 Pay As You Drive Service Execution

This component will classify the drivers' behaviour while driving according to the data collected from their car and exploiting the Driver Classyfier/Driving Profiling ML Models.

5.3.29 Investment Recommendation Engine Training

This component trains the personalized Investment Recommendation engine, to provide a set of recommendations for the financial instruments categories suitable for each customer and his/her investment profile, based on Market Index & Financial Instruments Sentiment data and Customer Risk Profiling ML Models. Machine Learning Models in Python mainly based on scikit-learn.

5.3.30 Investment Recommendation Engine Execution

Financial Instruments Personalized Investment Recommendation engine suitable for each customer and his/her investment risk profile, based on Market Index & Financial Instruments Sentiment data.

5.3.31 Recommender

Generates an actionable insight depending on the output from the remaining components within the Data Analytics layer, i. e. components P5b_Analytics_02 to P5b_Analytics_08.

5.3.32 Cash Flow Prediction

ML/AI model used to indicate and predict the available working capital (operating cash flow) of the SME (AS-IS & near term future).Alerts/notifications (via BOC Middleware PushNotifications) to be pushed to the respective SME in case of potential lack of liquidity and/or balance moving below a threshold.Cashflow Data/Insights to be provided to BOC Middleware Mobile/Web BPF.In order to provide valuable insights to the SME the data should be collected and streamed in real (or near real) time, since whenever a new object/ entry appears the model should retrain and adapt.

5.3.33 Budget Prediction

AI (ML) model is used to support the budget target setting for the various categories used by the respective SME. Doing so by providing budget predictions for each utilized category. The underlying model will take into consideration the cash flow analysis output, benchmark, macroeconomic and other available SME data (Business Plan).

5.3.34 KPI Engine

The KPI engine calculates KPIs with regards to the Financial Health and Performance of the respective SME.Doing so by taking into consideration the respective SME profile (e.g. maturity stage), accounting-wise optimal KPI values and how other similar SMEs perform.

5.3.35 Transaction (Txn) Monitoring

A dynamic complex event processing (CEP) mechanism that monitors the transactions of the user. In case transaction amount or type deviates from normal behaviour the user will be informed of abnormal transactions in order to be safeguarded from double payment mistakes and potential fraud attempts.In addition, expense pattern are also analysed to identify potential savings for instance multiple subscription spending or high ATM fees.

5.3.36 Transaction (Txn) Categorization

Smart transaction auto-classification which would also allow the user to manually override the given transaction category and define a new one (re-classify). The categorization performed will be based on the needs of the individual SME.

5.3.37 Invoice Processing

Processes transaction data and ERP data in order to keep record of the invoices which have been partially or fully paid by the SME. Provides insights regarding the respective VAT amount payable and the optimization of cash flow by providing background info on invoices such as paying a invoice at the “right” time.In cases where available invoice data is limited the engine will utilize a simplified approach in order to derive the expected VAT amount to be paid by the SME at the next VAT due date. Doing so by utilizing the banks transaction info or/and past VAT payment amounts.

5.3.38 KMeans

Batch BDA Service: Given a set of observations (x1, x2, …, xn), where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k ($\leq$n) sets S = {S1, S2, …, Sk} so as to minimize the within-cluster variance. It is a BDA service registered in the ALIDA catalogue. It is a PySpark-based micro-service running on K8S Spark cluster mode, working as part of the ALIDA framework.

5.3.39 Random Forest (Model)

Batch BDA Service: An ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set. It is a BDA service registered in the ALIDA catalogue. It is a PySpark-based micro-service running on K8S Spark cluster mode, working as part of the ALIDA framework.

5.3.40 Random Forest (Predict)

Streaming random forests algorithm. It will be a BDA service registered in the ALIDA catalogue. It is a PySpark-based micro-service running on K8S Spark cluster mode, working as part of the ALIDA framework.

5.3.41 Client Contextual Information

The component generates and updates a relevant client contextual information related to the clients' data and behavior. Open source AI/ML frameworks.

5.3.42 Financial Fraud/Crime Risk Score

Using the clients' contextual information and transactions a risk score of a fraudulent request of an online instant loan is evaluated.

5.3.43 Anomaly Analysis

Anomaly Analysis provides two main functionalities: • Anomaly detection • Anomaly prediction for time series data.

5.3.44 Pattern Analysis

Pattern Analysis provides two main functionalities:· Pattern matching· Discovery-The component will provide support for detection of complex patterns on data graphs.

5.3.45 Stream Story

Stream Story is a component for the analysis of multivariate time series. It computes and visualizes a hierarchical Markov chain model which captures the qualitative behaviour of the systems' dynamics, where system is described with a group of timeseries.

5.3.46 Open API Gateway

This component provides the single point of entry for the added-value functionalities (such as the Machine Learning (ML)/Deep Learning (DL) analytics functionalities) of INFINITECH which are based on microservices. The specific component enables the discovery and invocation of the dynamically registered microservices, effectively handling the incoming requests towards these microservice instances.

5.3.47 User Interface for Blockchain Transaction Reports and Visualization Component

This component will be responsible for providing user interaction with the Scalable Transaction Graph Analysis component within the bank and collect/manage user as well as annotated blacklisted blockchain addresses . It will utilize OpenAPIs (REST APIs) to submit queries consisting of customer blockchain addresses and blacklists to transaction graph analysis component and generate web based reports and visualization based received results. OpenAPIs/REST APIs, Web servers, Javascript, Vis.js graph drawing library (community version) (https://github.com/visjs-community/visjs-network).

5.3.48 Visualization Preparation

Stream BDA Service: Service to prepare data to the visualization depending on the type of incoming data or the data you want to view. It will be a microservice belonging to the ALIDA core.

5.3.49 Real Time Visualization

Useful tool for displaying charts through web application. Microservices deployed through Kubernetes • Docker • Data Source Connectors.

5.3.50 INFINISTORE

This is a GENERIC DATA STORE implementation for the INFINITECH Project as a microservice on top of a noSQL DB (MongoDB) instance. It is fed by different data ingestions servers and supports all other services. Microservice wrap on top of MongoDB. The microservice is implemented as a python-flask web server application on top of MongoDB instance. The API are defined with SWAGGER Open Api 3.0.

5.3.51 UI Risk Assessment Based on VaR

Web application to monitor portfolio risk in real time, perform what-if analysis, providing also several statistics of the underlying financial assets.

5.3.52 Pseudo-anonymization Tool

A tool that pseudo-anonymize data in order to preserve privacy. The component needs a specific configuration/development for each pilot in which it is used.

5.3.53 Health Insurance Risk Assessment Service

Algorithm yielding risk based on user RWD and a pre-trained classifier. Implemented as a Python script. Currently the assessment classifiers are: Random Forests or Logistic regression (using scikit- learn for inference), or neural networks (using Tensorflow for inference).

5.3.54 Health Insurance Fraud Detection Service

Algorithm detecting fraudulent behavior of insourance company customers. Implemented as a Python script.

5.3.55 Well-being Outlook Classifiers

Classifiers to be used by the health insurance risk assessment algorithm. The current set includes Random Forest, Logistic Regression and Neural Network classifiers of varying complexity. The format of the classifier depends on its type.

5.3.56 Synthetic RWD for Well-being Analytics

Synthetic data of the same format as those collected in Pilot 12. The current version of the data span 1,000 people simulated for 116 weeks each.

5.3.57 Open Banking Agreggator Solution

Crowdpolicy Open Banking Agreggator Solution is a modular architecture (UIs, connectors & APIs) platform so that it can be integrated into web/mobile banking applications, by the existing provider of the Bank in the form of API integration, but also as a separate application that could be made available to the users of Bank's online services Compatibility with best known market standards based on the European PSD2 Directive (Berlin Group, Open Banking UK, STET) Support for PISP & AISP services based on the PSD2 European Directive: – Payment Initiation Services – Account information Services.

5.3.58 Big Data Analytics Platform

Platform that collect and process information from multiple open data sources regarding SMEs and apply congnitive algorithm to detect risk and changes in financial needs. The tool will be use in pilot 13.

5.4 Processes

5.4.1 Semantics Streams Analytics Engine (SeSA-ME)

The INFINITECH Semantics Streams Analytics Engine (SeSA-ME) and the related tools for enabling semantic data exchange is based on the development of an interoperability (ontology-based) database/registry supporting linking of diverse systems and datasets based on shared semantics, as well as semantically interoperable analytics. The SeSA-ME system includes tools along with a visual SPARQL query editor providing Swagger APIs for verification and visualization tools for novice users while supporting full access and control over the data mashups for expert users. Tied with the development of the SeSA-ME platform is the development and deployment of the INFINITECH Graph Data Model which enables the support for both the design and deployment of stream-based web applications in a very simple and intuitive way and the analytics services using stream-based applications and services. [D4.6]

Semantic Streams Analyzer Middleware-Engine – SeSAME, The SeSAME component is a data mashup builder for the financial sector that can be used as a data processing component for your data management application, it enhances the capacity to process financial and insurance data in the form of batches and provides a single output, it is ideal when multiple sources have different data formats, it is built to be compatible with the most common data formats in the financial and Insurance sector i.e. FIBO, FIGI and LKIF and additionally it uses INFINITECH Core Graph Data Model to enhance performance. The SeSAME component is designed as a dataflow/workflow execution framework connecting various data input/outputs through the concept of pipelines for creating the data mashup. Conceptually, each financial operator has input data or streams and SeSAME provides one output data or stream. The multiple inputs can be used simultaneously while a single output in RDF is provided. Only the final operator of a workflow can return a format other than RDF, if necessary by defining and transforming the data into the desired format. The data Operators can be of three modes via APIs: API (1) a data acquisition operator is used to collect or receive data from data sources or gateways and can be pull-based or push-based. API (2) a stream processing operator defines stream processing functionalities in a declarative language, e.g., CQELS. API (3) a streaming operator streams the outputs of the final operator of a workflow to the consuming applications. In these three API modes operators of the data transformations and alignments can be done to produce a normalised RDF output format. [D4.4 to D4.6]

Enhanced Distributed Reasoner over FinTech Ontologies – EnDoRFIN Semantic Reasoner, The EnDoRFIN component is a tool for inferring knowledge from data streams, it uses some rules as conditions for defining logics conditions and as a result logical consequences are provided as outcomes. The inference rules are defined based on the most commonly used financial and insurance vocabularies i.e. FIBO, FIGI and LKIF and the way to process the rules is using APIs for defining the logical descriptions for the data applications it is introduced. This component allow the use of other languages but need to be upgraded to the target vocabulary and additionally the EnDoRFIN uses INFINITECH Core Graph Data Model to ensure the inference is applicable to all the involved domains in FINTECHs. [D4-4 to D4.6]

5.4.2 CI/CD

The “INFINITECH Way Foundations” summarizes ways in which partners of the project organize artefacts in the project’s code repository, and make use of the CI/CD pipelines to i) make their solutions available to other partners and ii) automate the deployment of a pilot. [Blueprint guidelines for the INFINITECH Way Foundations deployments of project pilots and technologies]

It makes use of an imaginary pilot solution that has been implemented to validate the process and can be used as a reference for other pilots to help them build their solutions. This reference pilot makes use of tow INFINITECH building blocks, the infinistore and the lx-kafka, and a micro-service that has been implemented specific for the needs of that pilot. In this document, we will see how we organize the code and artefacts: platform building blocks which are artefacts provided by INFINITECH, must go under the corresponding groups in order to be available for all pilot solutions, while pilot specific solutions must go under separate groups. We provide guidelines on how to create new projects under specific groups, setup the Dockerfile that instructs the CI process on how to build the image, and how to define the CI process itself, by making use of a Jenkins file. Finally, at the last step, we provide guidelines on how to make use of the CD process to define the artefacts that the solution consists of and how to automate the deployments. Blueprint guidelines for the INFINITECH Way Foundations deployments of project pilots and technologies.

INFINITECH Way Foundations can be used to setup the Dockerfile that instructs the CI process on how to build the image, and how to define the CI process itself, by making use of a Jenkins file. Finally, at the last step, it provides guidelines on how to make use of the CD process to define the artefacts that the solution consists of and how to automate the deployments. On the contrary with the CI pipelines, the CD pipelines are not triggered automatically. The reason is that we cannot afford to deploy all 16 Pilots solutions in the AWS resources. Due to this, the deployment will be started manually, for validation purposes. Blueprint guidelines for the INFINITECH Way Foundations deployments of project pilots and technologies.

5.4.3 KYC/KYB

The Know Your Customer (KYC)/Know Your Business (KYB) policies in state-of-the-art financial relations expect that customer parties, either individuals or entire corporations, endeavour verification of their identity. Thus, each financial organization is able to estimate the risks involved with sustaining a new business-customer partnership. In this context, together with the wider Finance field of Anti-Money Laundering (AML) procedures, every financial institution establishes KYC and KYB operations at the time they register a new customer. [D4.8]

The KYC/KYB blockchain application resolves the implementation of such industry mechanisms by utilizing blockchain technology as a basic background infrastructure. Security, immutability and controlled transparency rule inside large enterprise blockchain networks while they offer efficiency of tasks and effectiveness of transactions within the corporation and its members. Ultimately, blockchain technology simplifies the emerging use cases and directly addresses any possible issues that are created. [D4.7 to D4.9]

Particularly, a KYC/KYB mechanism ensures that the identification and verification of a customer occurs against national and international regulations and laws set by governments, commissions, central banks and financial associations. As both the customer profile information and the relevant laws and rules are subject to changes over time, their update and maintenance become complicated. Moreover, their centralized systems are exposed to data protection and cyber-security risks, which become cheaper to launch while they are led by more sophisticated adversaries year by year. [D4.7 to D4.9]

Blockchain technology and particularly permissioned blockchain networks are capable of providing security to the KYC and KYB processes through decentralization. The concept of decentralization mainly exploits the idea that the information is replicated across all network nodes, while sabotaging one or more nodes cannot harm the information integrity and a single point of failure is avoided. In particular, the permissioned blockchain technology promises to keep that sensitive information inside a private network where only privileged parties can access it with an insider invitation. Thus, the customer information is kept safe on a private ledger that offers transparency to a privileged group of legal network participants. Both the customer and the organization are able to perform create, read, write, delete (CRUD) operations on the data under pre-defined access control policies. The various features of permissioned blockchains enable different policies applications that are able to, for instance, separate legal parties into a higher privacy network running inside the initial private one. Improved privacy control and data immutability rule inside the aforementioned technological scenario while they ensure legitimate customer data protection and management together with proper administration of this data by financial enterprises. [D4.7 to D4.9]

Figure 5.2 depicts High-level architecture of the KYC/KYB solution. The KYC/KYB system is related to different stakeholders including customers, financial organizations, and financial institutions. It is ultimately connected with personalized blockchain network.

5.5 Self-assessment

The INFINITECH Readiness Level (IRL) is a self-assessment method created to evaluate and provide guidance during the adoption process of generic or specific technology assets and/or technology conditions related to the deployment phase of a pilot project. IRL‘s main objective is to act as a tool for project pilot leaders to help them and guide them in their process to identify and select ready-to-go technologies, in particular IRL also is useful in the on-boarding process of the INFINITECH Way Foundations where technologies that have been developed in advance can be adopted at any stage duration the execution of the project pilots. [Innovation Readiness Assessment]

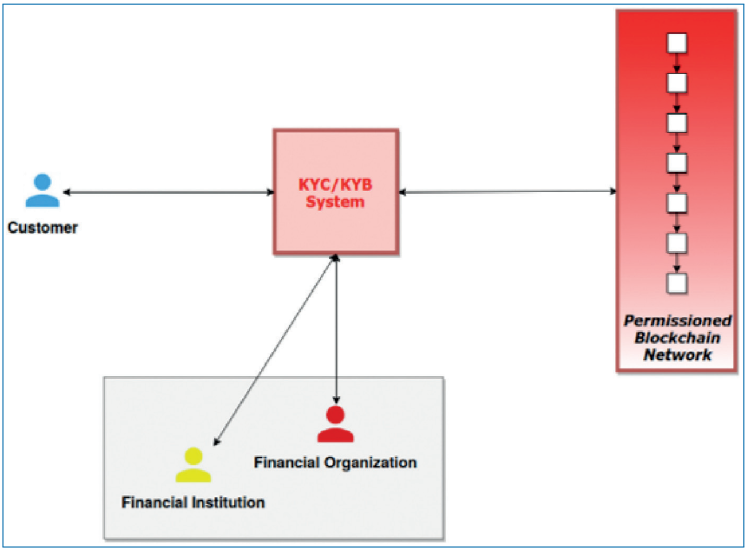


Figure 5.2. High-level architecture of the KYC/KYB solution.

IRL is applicable to any program (i.e. INFINITECH pilot) at any stage that operates or want to make use of the technology assets that are developed or implemented for the financial and insurance sectors, the IRL defines 5 levels scored 1 to 5 where 1 is the lowest level and 5 is the highest level. IRL can be used directly in the INFINITECH Pentagon (i.e. kiviat diagram) as it can be mapped directly within the innovation process/radar facilitating in this way the path towards achieve high levels of Innovation. [Innovation Readiness Assessment]

IRL provides to any project (INFINITECH pilot) the capacity to self-assess the level of adoption following identified technological characteristics in different technology domain areas i.e. Data Modeling and Data Interoperability, Infrastructure Deployment and Services Platform Adoption, Information Management and Analytics and Intelligent Applications. The technology domain areas are defined by following the requirements from ICT experts into a full-stack implementation and the IRL method allows the self-assessment of the evolution of the pilot project during the phase of adoption or onboarding process. IRL also adds the temporal dimension to an onboarding process which helps to not only self-assess the way the project is being implemented but also the validity of the process for following that path. [Innovation Readiness Assessment]

Figure 5.3 shows the INFINITECH Readiness Level (IRL) Self-Assessment Chart. We describe its components in below.

- IRL Level 1 – Look after the identification and definition of basic conditions to start a pilot with as much as possible re-use of INFINITECH baseline

Parameters	Level 1	Level 2	Level 3	Level 4	Level 5
Data Modeling	Vocabulary Identified	Taxonomy Ready	Data Model Logic & Physical	Validated Data Schema	Data Graph Ready
Data & Interoperability	Data Set Sample Ready	Data Storage Deployed	Query Data Tests Performed	Cross-Domain Query	Data Sharing/ Exchange
Trustworthiness Security & Privacy	Data Protection Methods	Access Control Tools	Identity Management	Platform Access Control	Self-Sovereign identity
Infrastructure	Local Host	Client-Server Mode	Cloud Environment	Docker Ready	Scale Up Tested
Services Platform	Communication Services using Service APIs	Management Services	Continuous Monitoring Tools	Services using Infinitech Orchestration Methods	Infinitech Orchestration and DevOps Compliance
Applications	Use Cases	Prototype	Demonstrator with videos	Online Services	Marketplace Ready

Figure 5.3. INFINITECH readiness level (IRL) self-assessment chart.

technologies. As a first step other technologies outside INFINITECH Ecosystem may be used but with a clear design perspective for on-boarding INFINITECH assets in next iteration.

- IRL Level 2 – Ready to go technologies for a prototype building in Client-Server model, a series of demonstrators can be available to test/proof the deployment of basic functionalities, at this level integration is not required but is recommended to have data model integration and common access control tools.
- IRL Level 3 – First Deployed Demonstrator with all functionalities in Cloud-based deployment, the use of INFINITECH Way Foundations and Reference Architecture is the core of this level, the use of demonstrators is essential to explain high level use cases and also the use of Data sets following the INFINITECH Data model or other best practices/standards in the particular domain of the pilot demonstration.
- IRL Level 4 – Cross Domain services deployed with DevOps ready for easy deployment. Data is required to be shared across different domains, it can be tested with simple applications or integrated in the cross-domain query system. At this level all the components are integrated following DevOps Techniques and the orchestration methods described in the INFINITECH Way Foundations.
- IRL Level 5 – Implemented Proof of Concept (PoC) with interoperable cross-domain services, DevOps and Sandboxes deployments are featured to facilitate the inclusion of the assets and solutions in the INFINITECH marketplace.

DOI: 10.1561/9781638282297.ch6

Chapter 6

INFINITECH Way Foundations Impact on Fintech and Insurance

6.1 INFINITECH Way Foundations Impact on Fintech and Insurance

The financial and insurance sector have not yet an adopted/accepted unified way of accessing & querying vast amounts of structured, unstructured, and semi-structured data. It is envisioning that a semantic approach can increase data interoperability and improve the OLTP (On-Line Transactional Processing) databases, OLAP (On-line Analytical Processing) databases and data warehouse which will reflect a potential benefit in the financial sector. The new technologies are having positive impact in all industries, and the FinTech's are not an exception, the effort and cost that is associated to finance and banking services with the development of BigData analytics and AI systems is compensated with the number of opportunities and economic benefits. [D4.4 to D4.7]

In recent years, the convergence of Internet technologies for communication, computation and storage networks and services has been a clear trend in the Information and Communications Technology (ICT) domain, beyond the fact that data fragmentation is an issue, there is also a lack of data interoperability across diverse datasets that can be reduced by using semantic technologies, however semantics can be used to alleviate this concurrent issue by using the semantic descriptions that refer to the same data entities with similar (yet different) semantics as the way to improve interoperability. Currently there is an increasing production of financial data and likewise an increase on the demand for such Information and in the other hand there is also a growing production of data coming from financial sectors, growing exponentially the number of sources of information, and thus it is necessary tools and systems that allows and facilitate that financial information can be accessed and integrated in a systematic, standardised, and cost-efficient manner. [D4.4 to D4.7]

Semantic web technologies are taking more relevance in the financial sector and systems where the information needs to be shared making the information readily useful for solving many scalability issues. Consequently, remarkable efforts have been invested to enable data interoperability, so that pieces of data can be plugged in into the data infrastructures, directly exposing their own data semantics instead of using the data itself, facilitating exchange services. By introducing semantic technologies, INFINITECH project provides an overlay that is much easier to process and at the same time minimize the risk on processing data. This semantic layer approach constitutes also the first step of the INFINITECH pipeline, i.e., gathering semantically annotated data from provided and/or available datasets or data streams. In this deliverable, we have described how INFINITECH project would benefit from semantic technologies like Linked Data and ontologies as the best practices in the semantic interoperability building process. [D4.4 to D4.7]

Following semantic best practices, we have design and implemented the Semantic Stream Analytics Middleware-Engine (SeSA-ME) analyzed the already existing ontologies that are related to the finance and insurance sectors and that can be reused for our purposes in the INFINITECH project. The main ontologies which are going to be used as baselines are FIBO, FIGI and LKIF, because they focused on both financial sector and financial operations containing the baseline for the metadata that represent, cross-domain and intra domain, financial transactions, and operations with an attached effort towards standardization. The INFINTECH Core ontology is an extension generated in the project that describes cross-domain vocabularies that are used in multi-domains within the INFINITECH project domain areas, it is meant to be complemented by other domain specific vocabularies. For this reason, and according to the initial requirements of the INFINITECH project, other vocabularies specifically related to security and payments are presented. [D4.4 to D4.7]

The project has successfully designed and used a holistic framework called INFINITECH Business Approach or "INFINITECH Way Foundations" to navigate through the design, development, and deployment of technology solutions in the financial services and insurance sectors.

The project develops and validates a rich set of novel BigData and AI systems including BigData architecture, Data store, Machine Learning (ML) algorithms, anonymization and digital on-boarding technologies, decentralized systems for digital finance, semantic interoperability technologies, etc.

In terms of foundation:

- The INFINITECH reference architecture (RA) has successfully been completed and shown how it applies to the 16 Pilots in development.
- The INFINISTORE database design has been completed providing a unified and integrated framework for handling real-time and offline data. It complements well the RA.
- Significant work has been completed with respect to data governance, handling, cleansing, pre-processing, and anonymization.
- A framework for semantic interoperability and data exchange across platforms, services and datasets.
- The library of ML algorithms has been extended, including a finalized version of a federated learning algorithm.
- The specification and prototype implementation of four blockchain prototypes for digital financial applications.
- Existing technologies (Flink, Calcite, Kafka, Etherum, Amazon WS, etc.) have been successfully leveraged to address the specific challenges faced in the banking and insurance sectors.

In terms of experimentation:

- Multiple sandboxes have been developed and implemented allowing to test and experiment with 16 Pilots and various assets made available through the INFINITECH marketplace.
- The virtual digital innovation hub (VDIH) has been integrated into the marketplace and effectively launched.
- The INFINITECH marketplace has been revised and the web presence updated, including numerous courses, workshops, and record webinars.

In terms of exploitation:

- First drafts of business models have been developed for each of the 16 Pilots (excluding the discontinued pilot 1) and the pilots evaluated by stakeholders relying on the INFINITECH Pentagram.
- Work has been performed in identifying synergies among consortium members by introducing the Joint Exploitation.

In term of framework

- For each of the pilots' lessons learned have been identified as a continuous quality improvement.

Also, the technology developments have an impact on the implicit training of different actors:

- Business users that have been training in the statistical nature of the results coming from such tools, and therefore the interaction and cooperation human-AI tools.
- Business managers to assess the impact and specificity of the use of such technologies.
- Data scientists and software architects and developers that have been trained during the development of the present project.

Frauds on financial services are an ever-increasing phenomena and cybercrime generates multi-million revenues, therefore even a small improvement in fraud detection rates would generate significant savings. This viewpoint, built on information sharing activities currently running in the banking sector, is also reinforced and strengthened by trusted industry reports.8910. With some surveys and reports pointing to issues, such as: "recover less than 25 percent of fraud losses", "Increase fraud typologies globally, from recent years, include identity theft and account takeover, cyber-attack, card not present fraud and authorized push payment scams", "6 is the average number of frauds reported per company studied", "56% asked companies conducted an investigation into their worst fraud incident. many organisations are failing to respond effectively". These, and other issues in these reports, demonstrate the importance of developing new technologies and approaches, such as real time analytics, to enhance the need of fighting against cyber frauds. [pilot #10]

Figure 6.1 illustrates the INFINITECH impact creation roadmap. It started with initial technical development and valuation. Followed by updated development and initial market valuation. Next final INFINITECH results, and market lunch was presented. Last, sustainability and INFINITECH wider use was kicked off.

Figure 6.2 shows the impact creation vehicles and dimensions. There are industrial, open source community, research and scientific, market and commercial, banks and financial institutions impacts. The impacts generated through INFINITECH solutions and technologies, INFINITECH VDIH, INFINITECH community and INFINITECH exploitation stories.

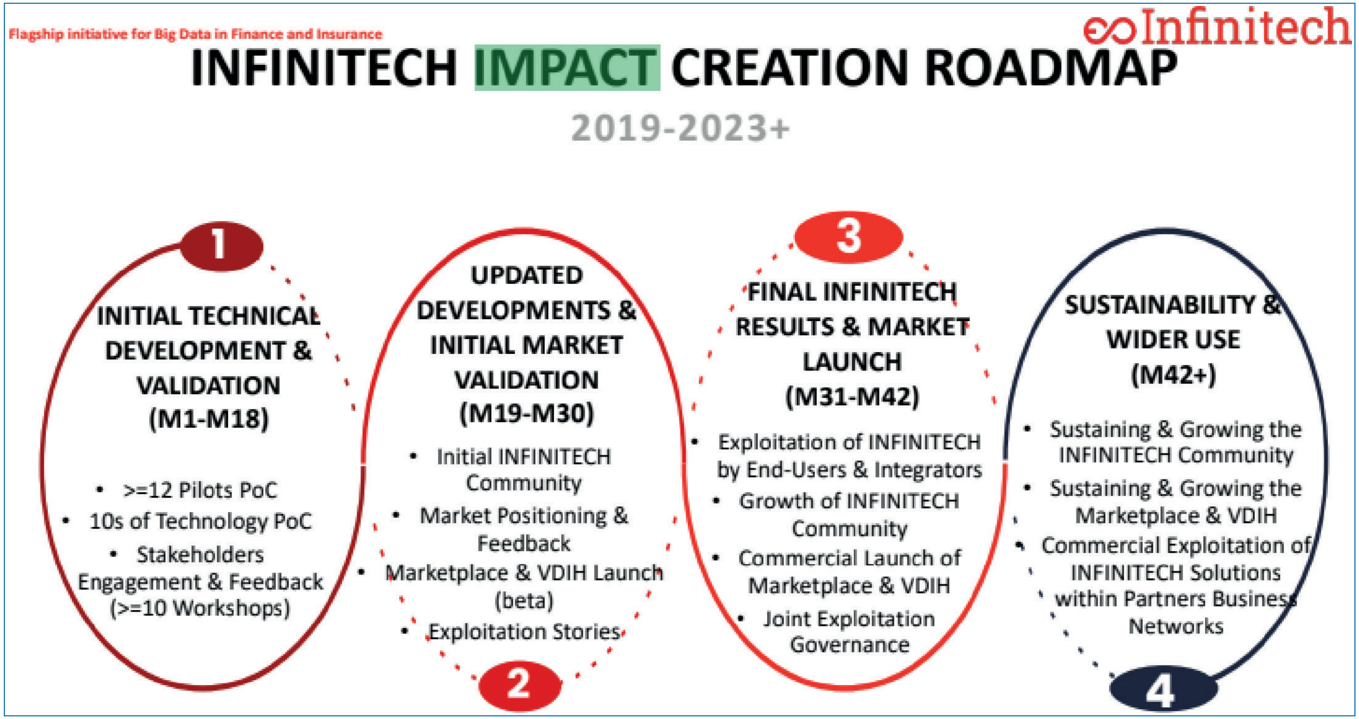
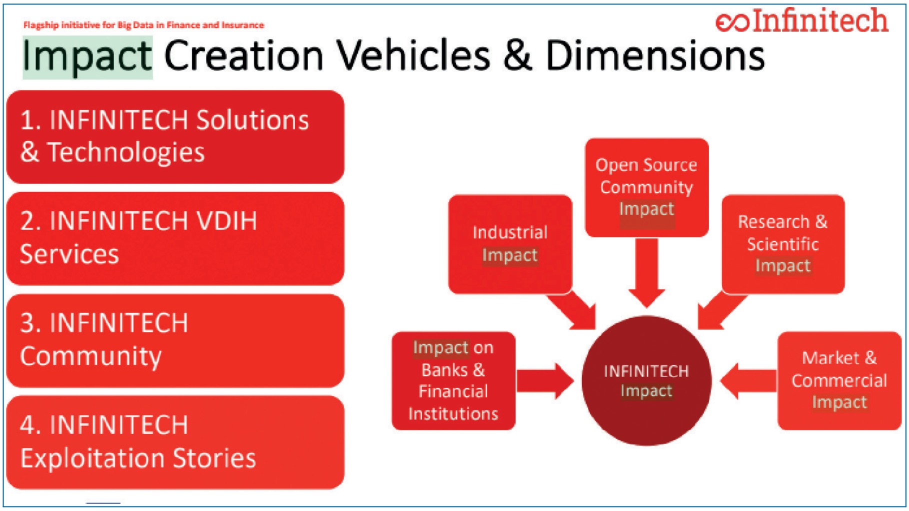

Flagship initiative for Big Data in Finance and Insurance
Infinitech
INFINITECH IMPACT CREATION ROADMAP
2019-2023+
1
INITIAL TECHNICAL DEVELOPMENT & VALIDATION (M1-M18)
• >=12 Pilots PoC
• 10s of Technology PoC
• Stakeholders Engagement & Feedback (>=10 Workshops)
2
UPDATED DEVELOPMENTS & INITIAL MARKET VALIDATION (M19-M30)
• Initial INFINITECH Community
• Market Positioning & Feedback
• Marketplace & VDIH Launch (beta)
• Exploitation Stories
3
FINAL INFINITECH RESULTS & MARKET LAUNCH (M31-M42)
• Exploitation of INFINITECH by End-Users & Integrators
• Growth of INFINITECH Community
• Commercial Launch of Marketplace & VDIH
• Joint Exploitation Governance
4
SUSTAINABILITY & WIDER USE (M42+)
• Sustaining & Growing the INFINITECH Community
• Sustaining & Growing the Marketplace & VDIH
• Commercial Exploitation of INFINITECH Solutions within Partners Business Networks

Figure 6.1 INFINITECH impact creation roadmap.



Flagship initiative for Big Data in Finance and Insurance

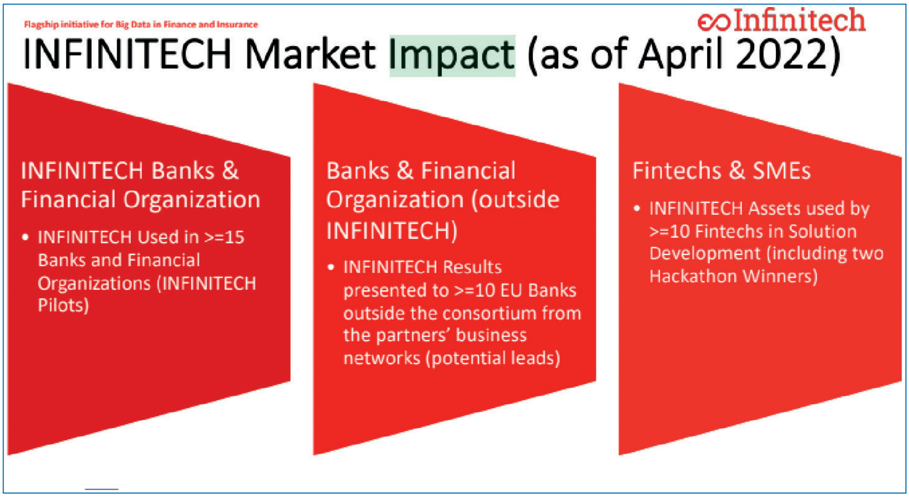
Infinitech

Impact Creation Vehicles & Dimensions

1. INFINITECH Solutions & Technologies
2. INFINITECH VDIH Services
3. INFINITECH Community
4. INFINITECH Exploitation Stories

Open Source Community Impact
Industrial Impact
Research & Scientific Impact
Impact on Banks & Financial Institutions
INFINITECH Impact
Market & Commercial Impact

Figure 6.2. Impact creation vehicles and dimensions.



Flagship initiative for Big Data in Finance and Insurance

Infinitech

INFINITECH Market Impact (as of April 2022)

INFINITECH Banks & Financial Organization

- INFINITECH Used in >=15 Banks and Financial Organizations (INFINITECH Pilots)

Banks & Financial Organization (outside INFINITECH)

- INFINITECH Results presented to >=10 EU Banks outside the consortium from the partners' business networks (potential leads)

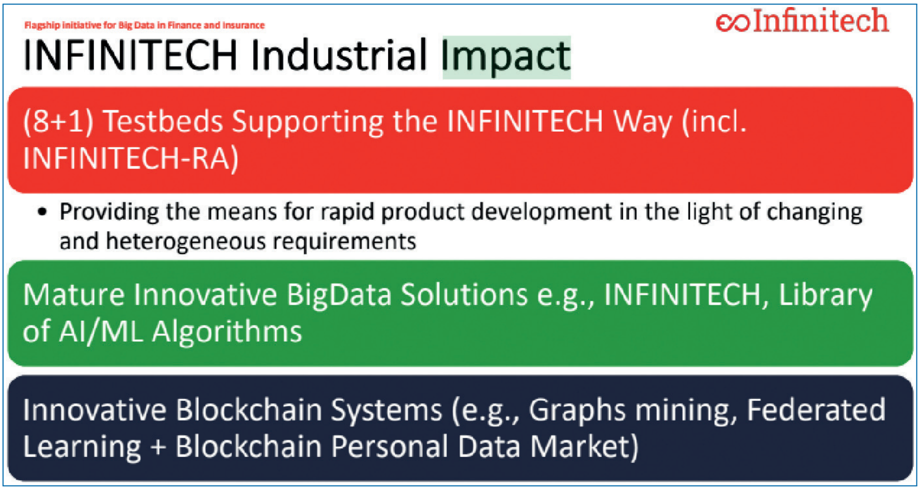
Fintechs & SMEs

- INFINITECH Assets used by >=10 Fintechs in Solution Development (including two Hackathon Winners)

Figure 6.3. INFINITECH market impact.

Figure 6.3 summarizes the INFNITECH market impact. INFINITECH has impact on banks and financial organization counterparts in INFINITECH project since it is used in more than fifteen banks and financial institutions (pilots). The INFINITECH project results were presented to more than ten banks and financial organizations outside INFINITECH. INFINITECH assets used by more than ten fintechs and small and medium size companies (SMEs).

INFINITECH has certain industrial impacts (as shown in Figure 6.4). There are nine testbeds that support the INFINITECH Way Foundations including



Flagship initiative for Big Data in Finance and Insurance

INFINITECH Industrial Impact

(8+1) Testbeds Supporting the INFINITECH Way (incl. INFINITECH-RA)

- Providing the means for rapid product development in the light of changing and heterogeneous requirements

Mature Innovative BigData Solutions e.g., INFINITECH, Library of AI/ML Algorithms

Innovative Blockchain Systems (e.g., Graphs mining, Federated Learning + Blockchain Personal Data Market)

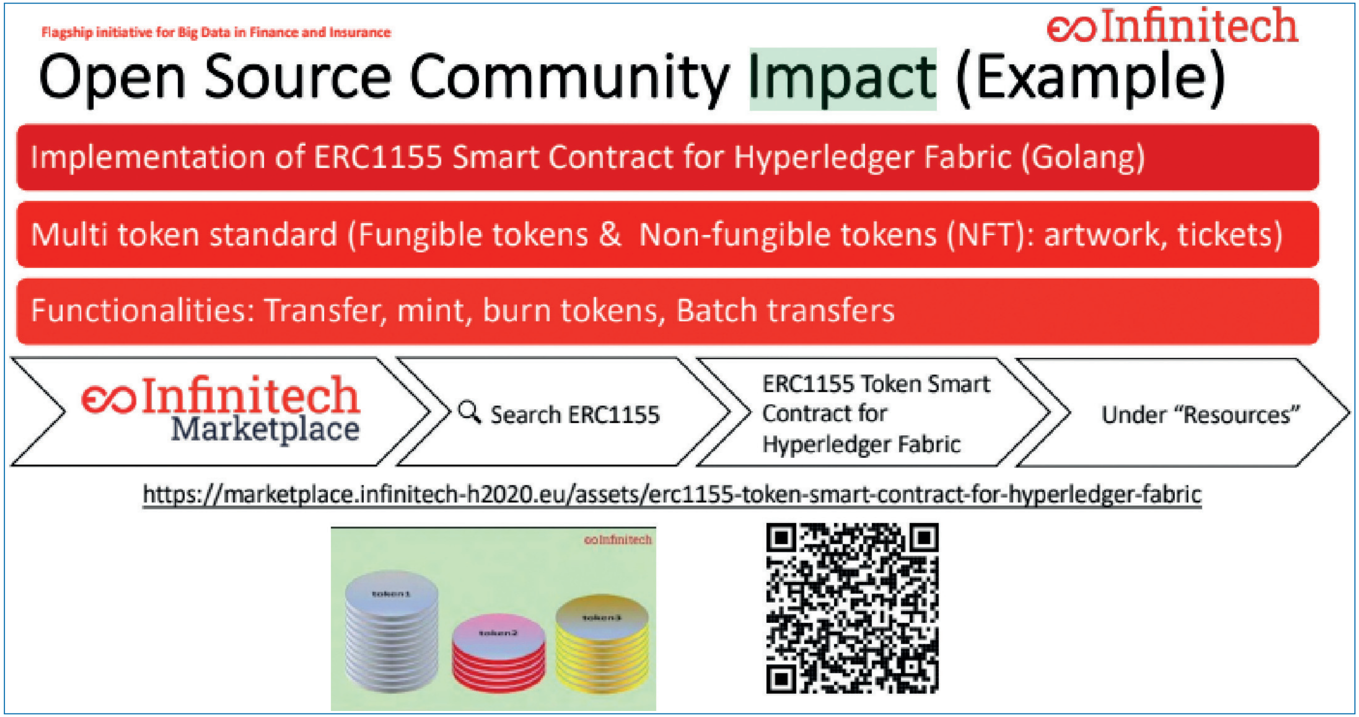
Figure 6.4. INFINITECH industrial impact.

INFINITECH-Reference Architect. These testbeds provide the means for rapid product development in the light of changing and heterogenous requirements. Further, INFINITECH developed mature innovative bigdata solutions such as INFINITECH, and library of AI/ML algorithms. The project devised innovative blockchain systems like graph mining, federated learning plus blockchain personal data market.

INFINITECH has open-source community impacts as well. It implements ERC1155 smart contract for Hyperledger fabric (Golang) (see Figure 6.5). It applies multi token standard including fungible token, non-fungible token, artworks, and tickets. It also introduces functionalities such as transfer, mint, burn token, and batch transfer.

INFINITECH has scientific impact as well (see Figure 6.6). The project produced an open access book. It is downloaded 17,000 times (in a week) an average of 1,000 per day. It is viewed almost 11,00 times on various Linkedin posts. It also includes additional impacts through scientific publications, conferences and open-source results. The scientific Impact also includes the three INFINITECH Series open access books.

Digitization is changing the Financial Services for many years from core banking to multichannel banking industry with different types of devices [1]. For several years now, the waves of digitization, financial technology (FinTech) and insurance technology (InsuranceTech) are rapidly transforming the financial and insurance services industry [3]. For instance, this is illustrated by the rapid growth of FinTech start-ups. McKinsey tracked more than 2000 FinTech start-ups in 2016 expecting even many more undetected [1]. Moreover, FinTech investments have grown from



Flagship initiative for Big Data in Finance and Insurance

∞Infinitech

Open Source Community Impact (Example)

Implementation of ERC1155 Smart Contract for Hyperledger Fabric (Golang)

Multi token standard (Fungible tokens & Non-fungible tokens (NFT): artwork, tickets)

Functionalities: Transfer, mint, burn tokens, Batch transfers

∞Infinitech Marketplace

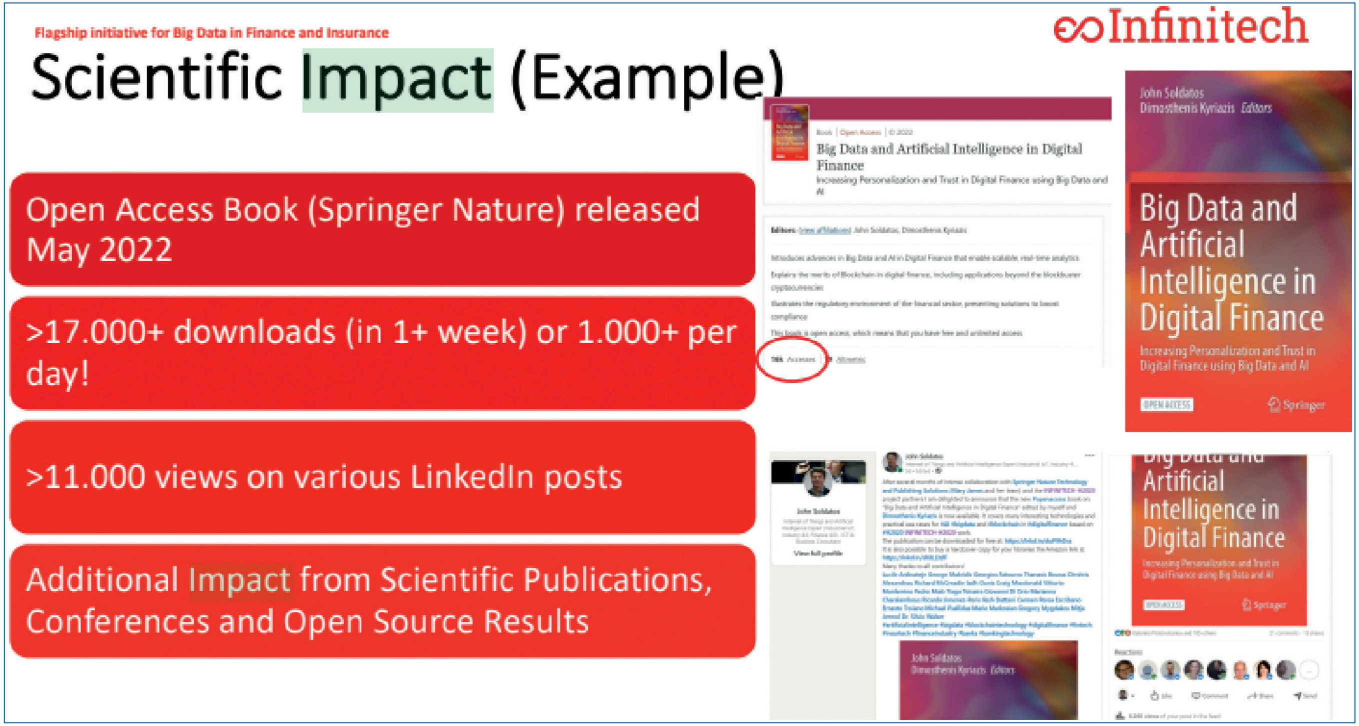
Search ERC1155

ERC1155 Token Smart Contract for Hyperledger Fabric

Under "Resources"

https://marketplace.infinitech-h2020.eu/assets/erc1155-token-smart-contract-for-hyperledger-fabric

Figure 6.5 Open-source community impact.



Flagship initiative for Big Data in Finance and Insurance

Scientific Impact (Example)

Open Access Book (Springer Nature) released May 2022

>17.000+ downloads (in 1+ week) or 1.000+ per day!

>11.000 views on various LinkedIn posts

Additional Impact from Scientific Publications, Conferences and Open Source Results

∞Infinitech
John Soldatos
Dimosthenis Kyriazis Editors
Big Data and Artificial Intelligence in Digital Finance
Increasing Personalization and Trust in Digital Finance using Big Data and AI
OPEN ACCESS
Springer
Big Data and Artificial Intelligence in Digital Finance
Increasing Personalization and Trust in Digital Finance using Big Data and AI
John Soldatos

Figure 6.6 Scientific impact.

1.8 bn USD in 2011 to more than 30 bn USD in 2018 with a CAGR of $\sim$ 50% p.a. [2]. [D2.3 snd D2.4]

FinTech is assigned to different views. They are as well attackers as enhancements of incumbents. Furthermore, FinTechs are part of large ecosystems, e.g. within the Alibaba platform, and selling infrastructure, e.g. used for open banking [2]. In challenging times or the end of a market cycle that is moving to a downward potential, new measures are required to maintain steady growth. Analysts like McKinsey [4] suggest several levers for organic growth to be explored by banks. [D2.3 and D2.4]

(1) risk management based on powerful analytical tools to prepare for a downturn;
(2) productivity, using modular utilities to materially change cost structures; and
(3) revenue growth through an improved customer experience (CX), bringing a larger customer base and/or share of wallet.

Essential to exploiting these profitability levers are the critical enablers of advanced data analytics and talent. AI shows a great promise in this field especially with the progress in modelling techniques and methods.

This will facilitate moving to new data sources as e.g. IoT supplementing traditional big data analytics in FinTech [2]. A rapid scaling of advanced analytics and AI tools is a key to successful growth from McKinsey's point of view [4]. For instance machine learning models can improve predictive accuracy in identifying the riskiest potential customers by up-to 35%. [D2.3 and D2.4]

According to Juniper Research [5] "*Technologies such as machine learning and blockchain are having a transformative effect on fintech, fundamentally altering the way financial services are delivered and driving fintech platforms to become the 'new normal'. Such technologies will make new use cases mainstream, including smart contracts, loan underwriting using AI to analyze non-traditional data sources, and personalized insurance policies based on IoT-generated data.*"

The highest impact of technology application in Financial and Insurance Services shall be obtained in the focus areas of INFINITECH [4, 5]:

- Lending & Financing, i.e. Credit Risks (Pilots 1 & 2).
- Wealth Management and Risk Assessment (Pilots 3 & 4).
- Customer Experience and Payment (Pilots 5 & 6).
- Regulation and Compliance, (Pilots 7 & 10).
- Usage based Insurance, which includes personalized insurance products (Pilots 11 to 14).
- Banking.

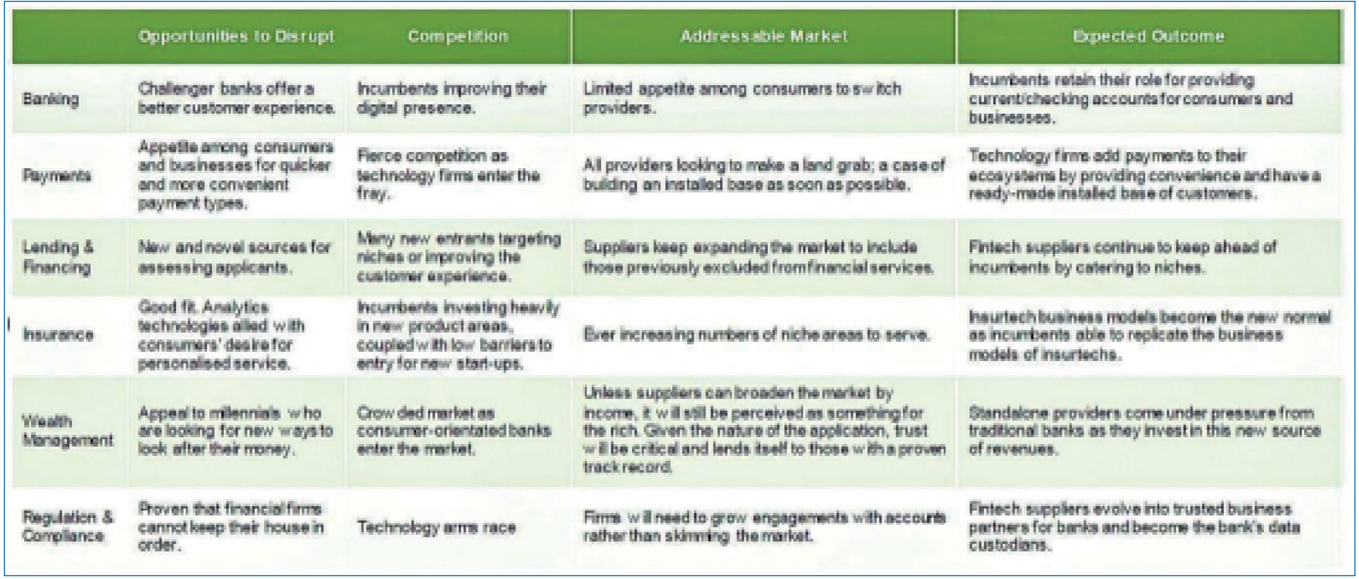
Figure 5.3 illustrates the INFINITECH impact in its focus areas. In banking area INFINITECH induce banks to offer a better customer experience. In this way, incumbents improve their digital presence. Customers have limited appetite to switch providers therefore, it is expected that incumbents remain the sole providers of current checking accounts for customers and businesses. In the area of payments there is a need for quicker and more convenient methods of payments among customers and business. This induces more competition among service providers, making providers looking to make a land grab as soon as possible. It is expected that technology firms add payments to their ecosystems by providing convenience and have ready-made installed base of customers.

In the lending and financing area, firms have new and novel sources for assessing applicants. Many new entrants target riches or improve the customer experience. Suppliers in this area keep expanding the market to include those previously excluded from financial services. FinTech suppliers continue to keep ahead of incumbents by catering to niches. Insurance companies use good fit, analytics, and technologies allied with consumers' desire for personalized services. Incumbent insurers invest heavily in new product areas coupled with low barriers to entry foe new start-ups.

There are ever increasing numbers of niche areas to serve. Nowadays, insurtech business models become the new normal as incumbents able to replicate business models of insurtechs. Wealth management is another area in which INFINITECH has major impact. It appeals to millennials who are looking for new ways to look after their investment. It is a crowded market as consumer-oriented banks enter the market.

The market in its current state is considered for rich that is the reason why supplier should consider broadening the market by income. Trust remains the most important issue in this market. As traditional banks begin to invest in this market, the standalone firms feel the pressure in their everyday business. Moreover, Regulation and Compliance, e.g. Financial Crime, Money Laundering, Fraud, includes strong opportunities to disrupt as Figure 6.7 shows. This area is covered by pilots 7 to 10.

As a side effect of the digital transformation in Financial Services, the trend towards persistent digital identities is accelerating. Indeed, "*this is due to multiple points of failure in conventional identification and verification processes, particularly for online payment details but also in a variety of other sectors. Passwords and centralized repositories have both been highlighted as the core issue within the growing problem of identity fraud, and a variety of approaches have arisen to combat this.*" [6]



	Opportunities to Disrupt	Competition	Addressable Market	Expected Outcome
Banking	Challenger banks offer a better customer experience.	Incumbents improving their digital presence.	Limited appetite among consumers to switch providers.	Incumbents retain their role for providing current/checking accounts for consumers and businesses.
Payments	Appetite among consumers and businesses for quicker and more convenient payment types.	Fierce competition as technology firms enter the fray.	All providers looking to make a land grab; a case of building an installed base as soon as possible.	Technology firms add payments to their ecosystems by providing convenience and have a ready-made installed base of customers.
Lending & Financing	New and novel sources for assessing applicants.	Many new entrants targeting niches or improving the customer experience.	Suppliers keep expanding the market to include those previously excluded from financial services.	Fintech suppliers continue to keep ahead of incumbents by catering to niches.
Insurance	Good fit. Analytics technologies allied with consumers' desire for personalised service.	Incumbents investing heavily in new product areas, coupled with low barriers to entry for new start-ups.	Ever increasing numbers of niche areas to serve.	Insurtech business models become the new normal as incumbents able to replicate the business models of insurtechs.
Wealth Management	Appeal to millennials who are looking for new ways to look after their money.	Crowded market as consumer-orientated banks enter the market.	Unless suppliers can broaden the market by income, it will still be perceived as something for the rich. Given the nature of the application, trust will be critical and lends itself to those with a proven track record.	Standalone providers come under pressure from traditional banks as they invest in this new source of revenues.
Regulation & Compliance	Proven that financial firms cannot keep their house in order.	Technology arms race	Firms will need to grow engagements with accounts rather than skimming the market.	Fintech suppliers evolve into trusted business partners for banks and become the bank's data custodians.

Figure 6.7 Fintech impact on financial services.

6.2 The Process of Adopting INFINITECH Way Foundations

The INFINITECH Way Foundations are adopted by 16 pilots in INFINITECH project which are listed below.

1. Configurable and Personalized Insurance Products
    - Pilot #13: Configurable and Personalized Insurance Products for SMEs
    - Pilot #14: Big Data and IoT for the Agricultural Insurance Industry
2. Personalized Retail and Investment Banking Services
    - Pilot #3: Collaborative Customer-centric Data Analytics for Financial Services
    - Pilot #4: Personalized Portfolio Management – Mechanism for AI based Portfolio Construction
    - Pilot #5A: Smart and Personalized Pocket Assistant for Personal Financial Management
    - Pilot #5B: Business Financial Management (BFM) tools delivering a Smart Business Advise
    - Pilot #6: Personalized and Intelligent Investment Portfolio Management for Retail Customer
    - Pilot #15: Open Inter-banking Pilot3. Personalized Usage-Based Insurance Pilots
3. Personalized Usage Based Insurance Products
    - Pilot #11: Personalized insurance products based on IoT connected vehicles.
    - Pilot #12: Real World Data for novel Insurance products
4. Predictive Financial Crime and Fraud Detection Pilots
    - Pilot #7: Avoiding Financial Crime
    - Pilot #8: Platform for Anti Money Laundering Supervision (PAMLS)
    - Pilot #9: Analysing Blockchain Transaction Graphs for Fraudulent Activities
    - Pilot #10: Real-time cybersecurity analytics on financial transactions' data
    - Pilot #16: Data Analytics Platform to detect payments anomalies linked to money laundering events

5. Smart, Reliable and Accurate Risk and Scoring Assessment Pilots
   - Pilot #1: Invoices Processing Platform for a more Sustainable Banking Industry
   - Pilot #2: Real time risk assessment in Investment Banking

DOI: 10.1561/9781638282297.ch7

Chapter 7

Conclusions

7.1 Conclusions

This book provides an easy way to understand the design principles and the overall process of the INFINITECH project, it can be considered as a handbook of the INFINITECH Way. Whit ever-increasing changes in the landscape of finance and insurance sector, technologies towards more human-centric applications are being developed. With the advancement of technologies, fin-tech and Insurance-tech business models are being re-defined. We explain the INFINITECH Way in simple words, enhancing the state of the art by introducing the design principles and basic metrics. We demonstrate the practicality and applications of INFINITECH Way in multiple areas.

In this book it is described and explained the advantages of INFINITECH Way in the on-boarding process which provide speed and agility in optimizing the design and implementation of financial services reference architecture. The INFINITECH Way mitigates the problems arising from dealing with complex vendor-lock and/or proprietary infrastructure. Also, in addressing Big Data, IoT and AI applications for the financial and insurance sectors, we expand on how INFINITECH Way is applied to validate a reference implementation.

In this book it is described the core elements of INFINITECH Way, the INFINITECH Reference Architecture (RA). A reference architecture for BigData systems in digital finance can greatly facilitate stakeholders in structuring, designing, developing, deploying and operating BigData, AI and IoT solutions. It serves as a stakeholders' communication device, while at the same time providing a range of best practices that can accelerate the development and deployment of effective systems. This book has introduced the first version of such a reference architecture, namely the INFINITECH-RA. The latter adopt the concept and principles of data pipelines, which are in-line with the state of the art in BigData and Artificial Intelligence systems. It is also in-line with the principles of the reference architecture of the BDVA.

In practice, INFINITECH-RA extends and customizes the BDVA RA with constructs that permit its use for digital finance use cases. The INFINITECH-RA defines the structuring principles that drive the integration of the INFINITECH technical components and technologies.

Furthermore in this book it is explained how INFINITECH-RA can be used to support the development of common BigData/AI pipelines for digital finance applications. In this direction, the book has provided some examples of popular pipelines. This book also provides an initial mapping of most pilots of the project to the INFINITECH-RA, using INFINITECH components and technologies. The presented RA will serve as a basis for the implementation of the initial versions of the MVP (Minimum Viable Product) of the pilots. These implementations will allow the collection of stakeholders' feedback regarding the RA. Based on this feedback, the next versions/released of the INFINITECH-RA will be produced, including updates in the different views of the architecture. Furthermore, future releases of the INFINITECH-RA will consider a broader range of BigData, IoT and AI related capabilities developed in INFINITECH.

In this first book series it is described the methodology applied for INFINITECH-RA which is based on review and analysis, architecture design, and fine tuning and updates. Further, we discuss general AI/BigData challenges such as different type of data sources, accessing heterogeneous data sources, and running analytics over stale data, analyzing real-time data to respond to events and create alerts and notifications as they events occur, and the existence of a lock between operational and analytical operations. Besides, we address specific challenges for financial sector. siloed data and business operations, real time performance requirements, mobility, multiple channels management, automation, transparency are among many challenges that INFINITECH Way is tackling.

The summary of the INFINITECH data pack which is the set of files, schemas, and metadata model diagrams (Graphs) that represent the way the INFINITECH data is organized and structured is described in details and the references to the relevant ontologies available in the market are included. Furthermore, all relevant taxonomies and vocabularies from INFINITECH Core, Financial Industry Business Ontology (FIBO), Financial Instrument Global Identifier (FIGI) and the Legal Knowledge Interchange Format (LKIF) domains used in INFINITECH are delineated and cited for more reference. Moreover, an example of Traffic Analysis

Hub Ontology (TAHO) to further the case is provided. Finally a detailed list of INFINITECH technologies, data, and processes is included.

This book details The INFINITECH Readiness Level (IRL) which is a self-assessment method created to evaluate and provide guidance during the adoption process of generic or specific technology assets and/or technology conditions related to the deployment phase of a pilot project. IRL's main objective is to act as a tool for project pilot leaders to help them and guide them in their process to identify and select ready-to-go technologies, in particular IRL also is useful in the on-boarding process of the INFINITECH Way Foundations where technologies that have been developed in advance can be adopted at any stage duration the execution of the project pilots.

Furthermore, the INFINITECH Way Foundations impact on FinTech and Insurance are described. It is of singular importance to know that the financial and insurance sector have not yet an adopted/accepted unified way of accessing & querying vast amounts of structured, unstructured, and semi-structured data. In this way INFINITECH Way provide a practical solution to real-world problems. The impact of INFINITECH Way on several categories including foundation, experimentation, exploitation, and framework is sumarized.

Finally, the process of adopting INFINITECH Way Foundations is described. Currently the INFINITECH Way Foundations are adopted by 16 pilots in INFINITECH project which are listed below under five categories. Each pilot is a showcase of how NFINITECH Way could be adopted by practical projects in the field.

DOI: 10.1561/9781638282297.ch8

Chapter 8

References

8.1 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

8.2 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, FINSEC 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. Efficient 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].