

Data Space Best Practices for Data Interoperability in FinTechs



Martín Serrano, Edward Curry, Richards Walsh, Gavin Purtill,
John Soldatos, Maurizio Ferraris, and Ernesto Troiano

Abstract This chapter focuses on data interoperability best practices related to semantic technologies and data management systems. It introduces a particular view on how relevant data interoperability is achieved and its effects on developing technologies for the financial and insurance sectors. Financial technology (FinTech) and insurance technology (InsuranceTech) are rapidly developing and have created new business models and transformed the financial and insurance services industry in the last few years. The transformation is ongoing, and like many other domains, the vast amount of information available today known as Big Data, the data generated by IoT, and AI applications and also the technologies for data interoperability, which allows data nowadays to be reused, shared, and exchange, will have a strong influence. It is evident the entire financial sector is in a moment of new opportunities with a new vision for substantial growth. This book chapter analyzes the basis of data space design and discusses the best practices for data interoperability by introducing concepts and illustrating the way to understand how to enable the interoperability of information using a methodological approach to formalize and represent financial data by using semantic technologies and information models (knowledge engineering). This chapter provides a state-

M. Serrano (✉)

Insight SFI Research Centre for Data Analytics, University of Galway, Galway, Ireland
e-mail: martin.serrano@nuigalway.ie

E. Curry

Insight SFI Research Centre for Data Analytics, University of Galway, Galway, Ireland
e-mail: edward.curry@nuigalway.ie

R. Walsh · G. Purtill

Banking and Payments Federation of Ireland – BPF Ireland, Dublin, Ireland
e-mail: richard.walsh@bpf.ie; gavin.purtill@bpf.ie

J. Soldatos

Athens Information Technology – AIT Athens, Athens, Greece

M. Ferraris · E. Troiano

CU Innovation Unit - GFT Italia Srl., Milan, Italy
e-mail: maurizio.ferraris@gft.com; ernesto.troiano@gft.com

of-the-art offer called INFINITECH Way using the discussed best practices and explains how semantics for data interoperability are introduced as part of the FinTechs and InsuranceTech.

Keywords FinTechs · InsuranceTech · Interoperability · Data Spaces · Semantics · Knowledge Graph · Linked Data · Ontologies and services · IoT · AI · Big Data

1 Introduction

Digitization or digital transformation, financial technology (FinTech), and insurance technology (InsuranceTech) are rapidly transforming the financial and insurance services industry [1, 2]. Although it is evident the entire financial sector is in a moment of new opportunities and visible, tangible growth, it is also apparent that this transformation is motivated for the FinTech and InsuranceTech enterprises, which are heavily disrupting the traditional business models, and the volume of relevant investments is a proof: Over \$23 billion of venture capital and growth equity has been allocated to FinTech innovations during 2011–2014, while \$12.2 billion was deployed in 2014 alone [3]. Moreover, a recent McKinsey & Co study revealed that FinTech start-ups in 2016 exceeded 2,000, from approx. 800 in 2015 [4]. Furthermore, most global banks and investment firms have already planned to increase their FinTech/InsuranceTech investments to yield a 20% average return on their investments. Again, beyond FinTech/InsuranceTech, financial institutions and insurance organizations are heavily investing in their digital transformation as a means of improving the efficiency of their business processes and optimizing their decision making.

Traditionally the financial and insurance services sectors and particularly the banking sector have been quite resistant to technology disruption. This is no longer the case in the current trend of digitizing society and its services and applications. The waves of the digital economy and unified markets demand new paradigms to be designed, implemented, and deployed. The vast majority of services and applications that have been developed for the finance and insurance sectors are data-intensive. This transformation holds for applications in different areas such as retail banking, corporate banking, payments, investment banking, capital markets, insurance services, financial services security, and mail. These applications leverage very large datasets from legacy banking systems (e.g., customer accounts, customer transactions, investment portfolio data), which they combine with other data sources such as financial market data, regulatory datasets, social media data, real-time retail transactions, and more. Disruptive innovation in finance and insurance is already possible today, for example, with the advent of Internet-of-Things (IoT) devices and applications (e.g., Fitbits, smartphones, smart home devices), several FinTech/InsuranceTech applications can take advantage of contextual data associated with finance and insurance services to offer a better quality of service at a

more competitive cost (e.g., personalized healthcare insurance based on medical devices and improved car insurance based on connected car sensors). Furthermore, alternative data sources (e.g., social media and online news) provide opportunities for new, more automated, personalized, and accurate services. Moreover, recent advances in data storage and processing technologies (including advances in Artificial Intelligence (AI) and blockchain technologies) provide new opportunities for exploiting the above-listed massive datasets, and they are stimulating more investments in digital finance/insurance services [5].

Financial and insurance organizations can take advantage of Big Data, IoT, and AI technologies to improve the accuracy and cost-effectiveness of their services and the overall value they provide to their corporate and retail customers. Nevertheless, despite early data space deployments, there are still many challenges that have to be overcome before leveraging the full potential of Big Data/IoT/AI in the finance and insurance sectors, which could also act as a catalyst for attracting more investments and for significantly improving the competitiveness of enterprises in these sectors [6].

This book chapter analyzes the basis of data space design and best practices for data interoperability by introducing concepts and illustrating the way to understand how to enable the interoperability of information using a methodological approach to formalize and represent financial data by using semantic technologies and information models (knowledge engineering) [7]. This chapter also focuses on the role that semantic technologies like Linked Data and information interoperability provide for the support of financial and insurance industries in the process of digital transformation.

The organization of this chapter is as follows: Section 2 presents challenges in the data space domain in terms of interoperability in the financial and insurance sectors where information exchange occurs to support Big Data, IoT, and AI-enabled services creation and delivery. Section 3 introduces the best practices for the data exchange approach in developments in several parallel streams. These streams facilitate information interoperability and act as a baseline supporting the information interoperability approach. Section 4 introduces the INFINITECH Way, a design, implementation, and deployment methodology to support FinTech Data Spaces. Section 5 presents the current state of the art and motivations for using semantic technologies in convergence and interoperability. Section 6 describes scalable features about the management of Linked Data and its benefits when used in the financial and insurance sectors. Section 7 presents the summary, and finally, some relevant references used in this chapter are listed.

2 Challenges in Data Space Design

Many of the challenges present in current Data Spaces [8] and information management systems are generated by data sharing and exchange, both considered data interoperability problems. Persistent challenges blocking progress in data space design and deployment are as follows.

2.1 Data Fragmentation and Interoperability Barriers

Nowadays, most of the data collected and possessed by financial organizations reside in a wide array of “siloes” (i.e., fragmented) systems and databases, including operational systems and OLTP (online transaction processing) databases, OLAP (online analytical processing) databases and data warehouses, data lakes (e.g., Hadoop-based systems) with raw data (including alternative data like social media), and others. In this fragmented landscape, heavy analytical queries are usually performed over OLAP systems, which leads financial organizations to transfer data from OLTP, data lakes, and other systems to OLAP systems based on intrusive and expensive extract-transform-load (ETL) processes.

In several cases, ETLs consume 75–80% of the budget allocated to data analytics while being a setup for seamless interoperability across different data systems using up-to-date data. Beyond the lack of integrated OLTP and OLAP processes, financial/insurance organizations have no unified way of accessing and querying vast amounts of structured, unstructured, and semi-structured data (i.e., as part of SQL and NoSQL databases), which increase the effort and cost that are associated with the development of Big Data analytics and AI systems. Moreover, there is a lack of semantic interoperability across diverse datasets that refer to the same data entities with similar (yet different) semantics beyond data fragmentation. This is a setback to sharing datasets across various stakeholders and enabling more connected applications that span multiple systems across the financial supply chain.

2.2 Limitations for Cost-Effective Real-Time Analytics

Most of the existing applications operate over offline collections of large datasets based on ETL (extract-transform-load) operations and fail to fully exploit the potential of real-time analytics, which is a prerequisite for a transition from reactive decisions (e.g., what to do following the detection of fraud) to proactive and predictive ones (e.g., how to avoid an anticipated fraud incident). Also, state-of-the-art near-real-time applications tend to be expensive as they have to persist large amounts of data in memory. Moreover, existing engines for real-time analytics (e.g., state-of-the-art streaming engines with stateless parallelization) have limitations when it comes to executing complex data mining tasks such as AI (deep learning-based) algorithms.

2.3 Regulatory Barriers

Big Data and IoT deployments must respect a complex and volatile regulatory environment. In particular, they must adhere to a range of complex regulations (e.g.,

PSD2 (Second Payment Services Directive), MiFIDII/MiFIDR (Markets in Financial Instruments Directive), 4MLD (fourth EU Money Laundering Directive) for financial/insurance) while at the same time complying with general regulations such as the GDPR (General Data Protection Regulation) and the ePrivacy directive. To this end, several RegTech initiatives aim at establishing regulatory sandboxes (e.g., [9–11]), i.e., specialized environments, that facilitate Big Data/IoT experimentation through ensuring access and processing of data in line with applicable laws and regulations. Nevertheless, the development of regulatory sandboxes is in its infancy and only loosely connected to leading-edge Big Data/IoT/AI technologies.

2.4 Data Availability Barriers

To innovate with IoT and Big Data, financial and insurance organizations (including FinTech/InsuranceTech innovators) need access to experimentation yet realistic datasets (e.g., customer account and payments' datasets) that would allow them to test, validate, and benchmark data analytics algorithms. Unfortunately, such data are hardly available, as their creation requires complex anonymization processes or even tedious processes that can realistically simulate/synthesize them. Hence, innovators have no easy access to data for experimentation and testing of novel ideas [12]. Also, due to the fragmentation of Europe's FinTech/InsuranceTech ecosystems, there are no easy ways to share such resources across financial/insurance organizations and innovators.

2.5 Lack of a Blueprint Architectures for Big Data Applications

Given the existing limitations (e.g., data silos and lack of interoperability), financial organizations are creating ad hoc solutions for their problems at hand. They leverage one or more instances of popular data infrastructures such as data warehouses, data lakes, elastic data stores, and machine learning toolkits in various deployment configurations. However, they have no easy way to create, deploy, and operate such infrastructures through adhering to proven patterns and blueprints that will lower their integration, deployment, and operation efforts and costs.

2.6 No Validated Business Models

Big Data and IoT deployments in finance/insurance have, in several cases, demonstrated their merits on the accuracy, performance, and quality of the resulting

services (e.g., increased automation in business processes, improved risk assessment, faster transaction completion for end-users, better user experience). However, there is still a lack of concrete and validated business models that could drive monetization and tangible business benefits for these service improvements. Such business models could foster the rapid development and adoption of Big Data and IoT innovations, including emerging innovations that leverage real-time analytics and AI [13].

3 Best Practices for Data Space Design and Implementation

To address these challenges and leverage the full potential of Big Data (including AI) and IoT in finance/insurance, there is a need for developments in several parallel streams.

3.1 Technical/Technological Developments

At the technical/technological forefront, there is a need for Big Data architectures and toolkits tailored to the needs of data-intensive applications in the finance/insurance sector. These shall include several novel building blocks, including (1) infrastructures for handling arbitrarily large datasets from multiple fragmented sources in a unified and interoperable way; (2) semantic interoperability solutions for the financial/insurance supply chain; (3) novel techniques for real-time analytics and real-time AI; (4) advanced data analytics algorithms (including AI); (5) technologies and techniques for security and regulatory compliance, such as data encryption and anonymization technologies; (6) blueprint architectures for combining the above-listed building blocks with coherent and cost-effective solutions; and (7) open APIs that will facilitate innovators to produce and validate innovative solutions.

3.2 Development of Experimentation Infrastructures (Testbeds)

The development of Big Data, IoT, and AI-based innovations requires significant testing and validation efforts, such as testing for regulatory compliance and optimizing machine learning and deep learning data models. Therefore, there is a need for widely available experimentation infrastructures at the national and EU levels, which shall provide access to resources for application development and experimentation, such as datasets, regulatory sandboxes, libraries of ML (machine learning)/DL (deep learning) algorithms, Open (banking/finance) APIs,

and more. Furthermore, such experimentation infrastructures should be available in appropriate testbeds, based on deploying the above-listed technical building blocks in various configurations. The latter should support experimentation and testing of all types of Big Data/AI/IoT applications in the finance and insurance sectors, such as KYC (Know Your Customer) and KYB (Know Your Business), credit risk scoring, asset management recommendations, usage-based insurance applications, personalized portfolio management, automated payment applications, and many more.

3.3 Validation of Novel Business Models

To showcase and evaluate the tangible value of the above-listed technologies and testbeds, there is also a need for validating them in the scope of real-life business cases involving realistic business processes and applications for retail and corporate finance/insurance. The validation shall focus on novel business models, which essentially disrupt existing operations of financial organizations and deliver exceptional business benefits in terms of automation, personalization, cost-effectiveness, and intelligence.

4 The INFINITECH Way to Design/Support FinTech Data Spaces

INFINITECH is the largest joint effort of Europe's leaders in IT and finance/insurance sectors toward providing the technological capabilities, the experimentation facilities (testbeds and sandboxes), and the business models needed to enable European financial organizations, insurance enterprises, and FinTech/InsuranceTech innovators to fully leverage the benefits of Big Data, IoT, and AI technologies. The latter benefits include a shift toward autonomous (i.e., automated and intelligent) processes that are dynamically adaptable and personalized to end-user needs while complying with the sector's regulatory environment. Furthermore, INFINITECH brings together all the stakeholders involving NGOs with their members, financial institutions and insurance companies, research centers, large industry, and SMEs.

4.1 Technological Building Blocks for Big Data, IoT, and AI

INFINITECH looks at the finance and insurance sectors and provides multiple assets, including infrastructures, components and toolkits for seamless data

access and querying across multiple fragmented data sources, technologies for cost-effective real-time analytics, advanced analytics algorithms (including AI), technologies for Data Governance and regulatory compliance, technologies for trustful and secure data sharing over blockchain infrastructures, as well as handling of semantic data interoperability across stakeholders of the financial/insurance supply chain. Furthermore, INFINITECH also follows reference architecture (RA) approach for Big Data, IoT, and AI applications in the financial sector [14–17], whose aim is to serve as a blueprint for integrating, deploying, and operating Big Data and IoT infrastructures, including infrastructures that will leverage the above-listed building blocks. Furthermore, the reference architecture provides the means for integrating and deploying applications that take advantage of leading-edge technologies, including predictive analytics, different instances of AI (e.g., DL, chatbots), and blockchains.

4.2 Tailored Experimentation Infrastructures

INFINITECH provides the necessary mechanisms for creating tailored experimentation environments (i.e., testbeds and sandboxes) for different applications (e.g., sandboxes for fraud detection, credit risk assessment, personalized financial assistance) using flexible configurations of the testbed resources. Testbeds and sandboxes are used for different Big Data, IoT, and AI applications in the financial and insurance sectors, enabling innovators to access and share resources for testing, innovation, and experimentation, including previous datasets. The INFINITECH testbeds and sandboxes use the Open API standard for experimentation, and innovation is crucial. This facilitates the adoption and the extension of the designed, deployed, and tested solutions. ML/DL algorithms and regulatory compliance tools play a relevant role in the tailored experimentation testbeds and sandboxes.

INFINITECH uses this concept by deploying testbeds and sandboxes European-wide, thus demonstrating that it is possible to support the FinTech/InsuranceTech partners through experimentation testbeds. INFINITECH includes seven testbeds established at individual banks and one (EU-wide). The testbeds are made available to innovators' communities via the established innovation management structures of the project's partners and through a (virtualized) digital innovation hub (VDIH) set by the project as part of its exploitation strategy.

4.3 Large-Scale Innovative Pilots in Finance and Insurance

The use of a large ecosystem like INFINITECH for testing and validation will leverage both the technological developments of the project (including the INFINITECH reference architecture) and the testbeds/sandboxes to later deploy and implement as part of commercial solutions the novel and validated use cases. The pilot's target

in real-life environments is based on realistic datasets, i.e., either anonymized or synthetic datasets with pragmatic statistical properties. The pilots will span a wide array of areas covering the most prominent processes of the financial and insurance sectors, including KYC and customer-centric analytics, fraud detection and financial crime, credit risk assessment, risk assessment for capital management, personalized portfolio management, risk assessment in investment banking, personalized usage-based insurance, insurance product recommendations, and more. The pilots will demonstrate the added value of the project's technologies and testbeds while at the same time showcasing the project's disruptive impact on Europe's financial and insurance sectors.

4.4 Business Model Development and Validation

In the scope of the innovative pilots and use cases in finance and insurance, notably a novel and replicable business model or a set of them needs to be associated with each of the listed pilots/use cases [18, 19]. A practice to resolve one of the significant issues when developing new technologies based on experimentation is the use of a real exploitation model. These business models will pave the ground for disrupting the financial sector based on advanced Big Data, IoT, and AI infrastructures and applications, thus demonstrating the tangible impact of the project in financial institutions, insurance organizations, and FinTech/InsuranceTech enterprises.

5 Technology Capabilities for Convergence and Interoperability

Financial technology (FinTech) and insurance technology (InsuranceTech) are rapidly developing and have created new business models and transformed the financial and insurance services industry in the last few years. Technological convergence supporting data sharing and exchange between services applications is a barrier that the financial and insurance sectors have recently confronted with the globalization of economies and markets for a long time. This need is becoming more relevant, and today more than ever before, it needs to be addressed. Semantic technologies have played a crucial role as an enabler of many of the applications and services in other domain areas, although not much in the financial domain, and as has already been mentioned in the financial and insurance sectors, it is until just recently that the progress, in terms of implementation, has become more evident requirements; however, the convergence between technological development and interoperability has not entirely run in parallel, mainly due to many complex issues involving non-interoperable aspects where social, economic, and political dimensions are taking place.

5.1 *Semantic Interoperability and Analytics*

INFINITECH provides a shared semantics solution for the interoperability of diverse finance/insurance datasets. To this end, the project relies on existing ontologies for financial information modeling and representation (such as FIBO, FIGI, and LKIF) [20, 21], which are appropriately extended as required by the INFINITECH-RA and the project's pilots. Moreover, INFINITECH offers a solution for parallel and high-performance analytics over semantic streams, based on the customization of existing solution of semantic linked stream analytics (such as NUIG's Super Stream Collider (SSC) solution [22–25]). The INFINITECH semantic interoperability infrastructure is available in all cases/pilots where semantic reasoning will be required for extra intelligence.

5.2 *INFINITECH Building Blocks for Big Data, IoT, and AI*

There is always a high demand for integrated systems and technological components that can almost transparently connect and transfer data in the finance and insurance sectors. The integrated environment, including infrastructures, components, and toolkits, shall be designed to support seamless data access and querying across multiple fragmented data sources, technologies for cost-effective real-time analytics, advanced analytics algorithms (including AI), technologies for Data Governance and regulatory compliance, technologies for trustful and secure data sharing over blockchain infrastructures, as well as handling of semantic data interoperability across stakeholders of the financial/insurance supply chain. INFINITECH emerges as an alternative to those technological demands and provides a reference architecture (RA), as shown in Fig. 1. The INFINITECH reference architecture brings together technologies for Big Data, IoT, and AI applications in the financial sector, which will serve as a blueprint for integrating, deploying, and operating Big Data and IoT infrastructures.

INFINITECH provides the means for integrating and deploying applications that take advantage of leading-edge technologies, including predictive analytics, different instances of AI (e.g., DL, chatbots), and blockchains between other technologies. Figure 1 depicts the mentioned innovation-driven functional architecture approach from the INFINITECH ecosystem. It is an overall and FinTech holistic view. Its design and implementation rely on the intelligence plane, a combination of Big Data, IoT, and AI analytics applications. In the INFINITECH reference architecture data analytics plane, the exchange of information facilitates knowledge-driven support and the generation of composing services with operations by enabling interoperable management information.

The INFINITECH approach, in terms of the design approach, uses the design principles introduced in this chapter and looks at implementing the different scenarios and testbeds as described. INFINITECH moves toward converged IP

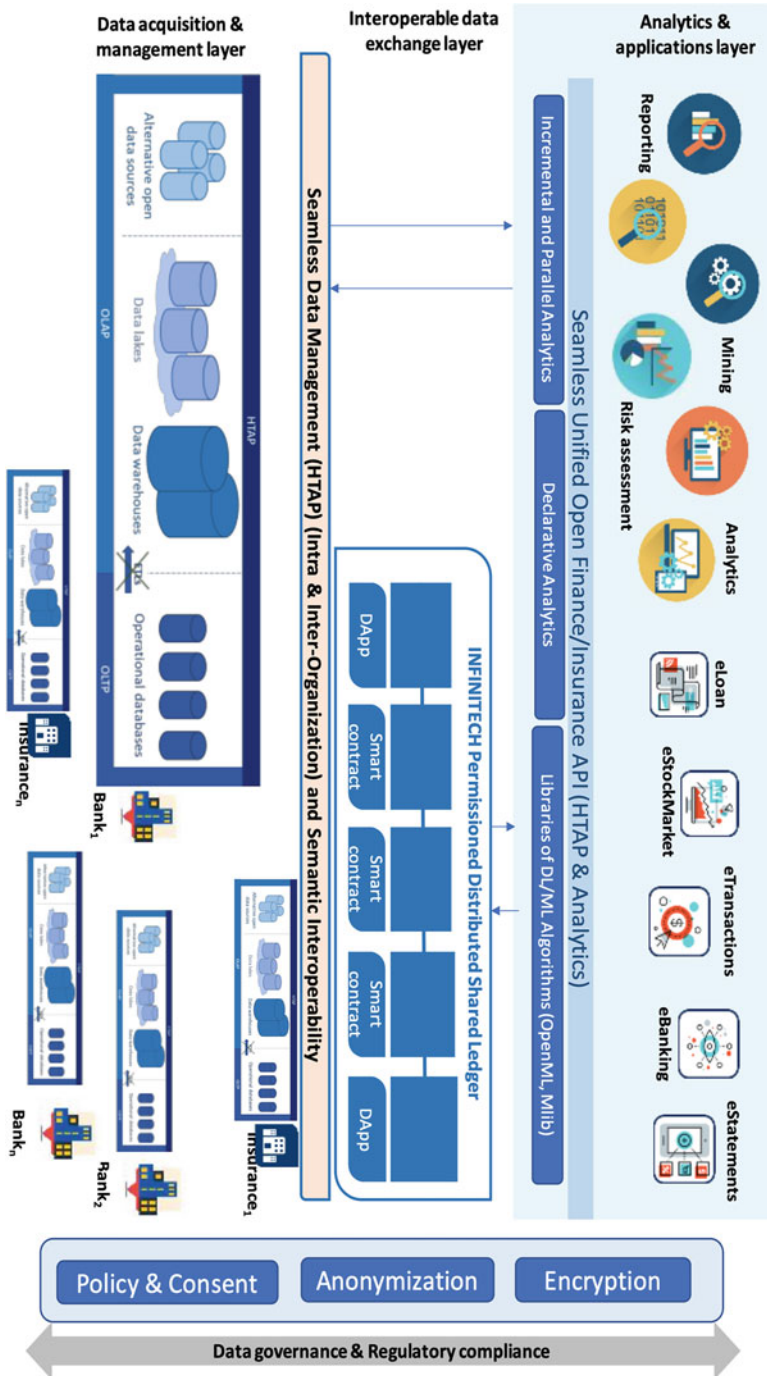


Fig. 1 INFINITECH reference architecture – high-level overview of Big Data/IoT platforms and technological building blocks

and cloud-based communications networks, increasing solutions to a number of significant technical issues by using more standard information exchange, promoting information interoperability, and allowing that the testbeds and sandboxes be managed effectively, and, most importantly, offering new open opportunities for a user knowledge-based service-oriented support can have a fundamental impact on future financial and insurance services.

6 Scalability and Security Considerations for FinTech and InsuranceTech

There are basic characteristics that must be taken into consideration at the time of building new approaches for the FinTech and InsuranceTech industries. Relatively modern ways to build multiple applications are sandboxes and testbeds specialized for providing close to real deployment, and thus implementations can be tested under real digital infrastructure conditions. The Big Data/IoT technologies and sandboxes/testbeds must be coupled with novel business models that will enable a whole new range of novel applications that will emphasize automation, intelligence, personalization, security, stakeholders' collaboration, and regulatory.

INFINITECH provides a 360° coverage of all the issues that hinder financial institutions' and FinTech enterprises' efforts to use and fully leverage IoT and Big Data technologies, including AI. Thus, this section summarizes pragmatically (following INFINITECH experiences) how those efforts, for scalability, are done.

Compliance and at the same time INFINITECH will enable the development, deployment, and business validation of a whole new range of applications that will be characterized by SHARP (Smart, Holistic, Autonomy, Personalized and Regulatory Compliance) characteristics. The following are just short paragraphs describing how SHARP can be implemented and briefly and as implementation reference how they have been addressed in INFINITECH.

- **Smart:** Services shall take advantage of predictive analytics and AI on Big Data datasets to anticipate changes in financial/insurance contexts and automatically adapt to them. INFINITECH has designed a set of end-to-end, business-to-customer, or business-to-business applications. Those services are based on analytics or ML solutions.
- **Holistic:** Architectures shall empower a holistic approach to data-driven services, which shall support all different financial applications across all phases of their lifecycle, including applications spanning multiple stakeholders and systems in the financial supply chain. INFINITECH implemented a series of reference architectures in the form of functional components.
- **Autonomy:** The deployed infrastructures shall take advantage of Big Data, IoT, and AI to significantly reduce error-prone manual processes and decision making through increased automation. The INFINITECH reference architecture presented paves the ground for the fully autonomous processes in the future that

will disrupt the entire finance and insurance sectors. The use of the architecture is a sign of autonomy, and it does contain all the necessary software parts to run and operate services.

- **Personalization:** Processes where data is processed timely profile customers and subsequently offer them individualized and dynamic services that adapt to their needs. INFINITECH involved KYC/KYB processes to include individuals' characteristics and build profiles based on the available data.
- **Regulatory Compliant:** Based on the use of particular or general technologies, the financial and insurance enterprises can ensure regulatory compliance by design. INFINITECH will take advantage of data processing to achieve faster, cost-effective, and reliable compliance to regulations.

To ensure scalability and security, permissioned blockchain for data sharing and data trading is required. There are already several use cases in the finance/insurance sector [26] that involve sharing of data across different organizations (e.g., sharing of customer data for customer protection or faster KYC, sharing of businesses' data for improved credit risk assessment, sharing of customer insurance data for faster claims management, and more); these are the ideal scenarios for emerging solutions like DLT (distributed ledger technologies, the baseline for blockchain technologies).

INFINITECH uses a permissioned DLT infrastructure, which provides privacy control, auditability, secure data sharing, and faster operations. Some of the later characteristics are inherent in permissioned DLT's features (i.e., the Hyperledger Fabric by IBM and the Linux Foundation) and can be directly configured in INFINITECH testbeds, sandboxes, and pilots. The core of the DLT infrastructure, i.e., the Fabric, will be enhanced in two directions: (1) integration of tokenization features and relevant cryptography as a means of enabling assets trading (e.g., personal data trading) through the platform and (2) for selected pilots enhancement of the blockchain infrastructure with multi-party computation (MPC) and linear secret sharing (LSS) algorithms from the OPAL (Open Algorithm Project) to enable querying of encrypted data as a means of offering higher data privacy guarantees.

7 Conclusions

This chapter addresses best practices for data space design and implementation identified from the state-of-the-art analysis. These challenges are tested and validated in the context of an H2020 European large-scale ecosystem called INFINITECH. The best practices leverage the full potential of Big Data, IoT, and AI applications in finance/insurance and identify a need for developments in several other areas to support scaling-up applications.

The interoperable data model following the formalization of vocabularies [7, 27] and using the FIBO, FIGI, and LKIF approaches were mentioned. It is the most adaptive interoperability model for the financial and insurance sector. Although the details are out of the scope of this chapter, the references and methods to use

them as part of the formal vocabularies to build the FinTech and InsuranceTech lingua franca are already an innovative approach toward data exchange and data sharing capabilities and are introduced as part of the results from the INFINITECH ecosystem approach called INFINITECH Way.

The INFINITECH reference architecture has been introduced, which provides the means for integrating and deploying applications that take advantage of leading-edge technologies, including predictive analytics, different instances of AI (e.g., DL, chatbots), blockchains, and more. In addition, INFINITECH provides technology for semantic interoperability based on shared semantics, along with a permissioned blockchain solution for data sharing across finance/insurance institutions.

This book chapter analyzed the basis of data space design and discussed the best practices for data interoperability by introducing concepts and illustrating the INFINITECH Way to enable interoperability of the information using a reference architecture following the methodological approach and the formalization and representation of financial data by using semantic technologies and information models (knowledge engineering). In addition, the INFINITECH Way introduced discusses best practices and explains how challenges for data interoperability can be overcome using a graph data modeling approach as part of the FinTech and InsuranceTech.

Acknowledgments This work has been partially supported by the H2020 INFINITECH Project (Tailored IoT & BigData Sandboxes and Testbeds for Smart, Autonomous and Personalized Services in the European Finance and Insurance Services Ecosystem) [28], which is co-funded by the European Commission under the H2020 framework program, contract number H2020-ICT-2019-856632-INFINITECH. It is also partially supported by Science Foundation Ireland under grant number SFI/12/RC/2289_2.

References

1. Dietz, M., Khanna, S., Olanrewaju, T., & Rajgopal, K. *Cutting through the noise around financial technology*. McKinsey & Co, February 2016, Accessible online Oct 2021, from <https://www.mckinsey.com/industries/financial-services/our-insights/cutting-through-the-noise-around-financial-technology>
2. PwC. *Redrawing the lines: FinTech's growing influence on Financial Services*. Global FinTech Report 2017, March 2017. Available at; <https://www.pwc.com/jg/en/publications/pwc-global-fintech-report-17.3.17-final.pdf>
3. Crofoot, J., Jacobs, P., Sarvady, G., Schultz, A. B., & Willder, J. *McKinsey on payments: Special edition on advanced analytics in banking*, April 2018.
4. McKinsey & Company. *Financial services practice - global payments 2015: A healthy industry confronts disruption*. Available Online Oct 2021 https://www.mckinsey.com/~media/mckinsey/dotcom/client_service/financial%20services/latest%20thinking/payments/global_payments_2015_a_healthy_industry_confronts_disruption.ashx
5. Brynjolfsson, E., & McAfee, R. The business of artificial intelligence. *Harvard Business Review (HBR)*, July 17.
6. Zillner, S., Curry, E., Metzger, A., Auer, S., & Seidl, R. (Eds.). (2017). *European big data value strategic research & innovation agenda*. Big Data Value Association. Retrieved from http://www.edwardcurry.org/publications/BDVA_SRIA_v4_Ed1.1.pdf

7. Serrano, J. M., Serrat, J., & Strassner, J. Ontology-based reasoning for supporting context-aware services on autonomic networks. [2007 IEEE/ICC International conference on communications](#), 24–28 June 2007, Glasgow.
8. Curry, E. (2020). *Real-time linked dataspace*s. Springer International Publishing. <https://doi.org/10.1007/978-3-030-29665-0>
9. *Industry SandBoxes*. Available on October 2021. <http://industrysandbox.org/regulatory-sandboxes/>
10. *Financial Conduct Authority*, “Regulatory sandbox”, White Paper, November 2015.
11. Fernandez, R. C., Migliavacca, M., Kalyvianaki, E., & Pietzuch, P. (2014). Making state explicit for imperative big data processing. In *USENIX ATC*.
12. Curry, E., Metzger, A., Zillner, S., Pazzaglia, J.-C., & García Robles, A. (2021). *The elements of big data value*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-68176-0>
13. Zillner, S., Bisset, D., Milano, M., Curry, E., Hahn, T., Lafrenz, R., et al. (2020). *Strategic research, innovation and deployment agenda - AI, data and robotics partnership. Third Release (Third)*. Brussels: BDVA, euRobotics, ELLIS, EurAI and CLAIRES.
14. Pääkkönen, P., & Pakkala, D. (2015). Reference architecture and classification of technologies, products and services for big data systems. *Big Data Res*, 2(4), 166–186. <https://doi.org/10.1016/j.bdr.2015.01.001>
15. Muan Sang, G., Xu, L., & De Vrieze, P. (2016). *A reference architecture for big data systems* (pp. 370–375). <https://doi.org/10.1109/SKIMA.2016.7916249>.
16. Muan Sang, G., Xu, L., & De Vrieze, P. (2017). *Simplifying big data analytics systems with a reference architecture* (pp. 242–249). doi: https://doi.org/10.1007/978-3-319-65151-4_23.
17. Schiff, A., & McCaffrey, M.. *Redesigning digital finance for big data*, May 2017.
18. Simitsis, A., Wilkinson, K., Castellanos, M., & Dayal, U. (2012) Optimizing analytic data flows for multiple execution engines. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, New York, NY, pp. 829–840.
19. Sun, Z., Strang, K. D., & Yearwood, J. (2014). Analytics service oriented architecture for enterprise information systems (pp. 508–516). doi:<https://doi.org/10.1145/2684200.2684358>.
20. Object Management Group. *Financial instrument global identifier specification v1.0*, December 2015.
21. Bennett, M. (2013). The financial industry business ontology: Best practice for big data. *Journal of Banking Regulation*, 14(3–4), 255–268.
22. Le-Phuoc, D., Nguyen-Mau, H. Q., Parreira, J. X., & Hauswirth, M. (2012). A middleware framework for scalable management of linked streams. *Journal of Web Semantics*, 16.
23. Neumeyer, L., Robbins, B., Nair, A., & Kesari, A. (2010). S4: Distributed stream computing platform. *IEEE International Conference on Data Mining Workshops, 2010*, 170–177.
24. Nguyen, H., Quoc, M., Serrano, M., Le-phuoc, D., & Hauswirth, M. (2012). Super stream collider linked stream mashups for everyone. In: *Proceedings of the Semantic Web Challenge co-located with ISWC2012*, vol. 1380.
25. Serrano, M., Strassner, J., & ÓFoghlú, M. A formal approach for the inference plane supporting integrated management tasks in the future internet. *1st IFIP/IEEE ManFI International Workshop, In conjunction with 11th IFIP/IEEEIM2009*, 1–5 June 2009, at Long Island, NY.
26. Carson, B., Romanelli, G., Walsh, P., & Zhumaev, A., *Blockchain beyond the hype: What is the strategic business value?*, June 2018.
27. Serrano, J. M. (2012). *Applied ontology engineering in cloud services, networks and management systems*. Springer, Hardcover, 222 pages, ISBN-10: 1461422353, ISBN-13: 978-1461422358.
28. H2020 INFINITECH Project. Accessible online October 2021. <http://www.infinitech-project.eu>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

