Tailored IoT & BigData Sandboxes and Testbeds for Smart, Autonomous and Personalized Services in the European Finance and Insurance Services Ecosystem

∞Infinitech

D8.2 – Market Platform and VDIH Specifications - II

Revision Number	3.0
Task Reference	T8.1
Lead Beneficiary	UPRC
Responsible	Dimosthenis Kyriazis, Vasilis Koukos
Partners	CPH, DYN, GFT, INSO, INNOV, UNP, UPRC
Deliverable Type	Report (R)
Dissemination Level	Public (PU)
Due Date	2021-03-31
Delivered Date	2021-03-31
Internal Reviewers	NOVA, AGRO
Quality Assurance	INNOV
Acceptance	Coordinator Accepted
EC Project Officer	Pierre-Paul Sondag
Programme	HORIZON 2020 - ICT-11-2018

Contributing Partners

Partner Acronym	Role[1]	Author(s)[2]
UPRC	Lead Beneficiary	Vasilis Koukos, Dimosthenis Kyriazis
GFT	Contributor	Vittorio Monferrino
INNOV	Contributor	John Soldatos
CP	Contributor	Marinos Xynarianos
UNP	Contributor	Tiago Teixeira, Bruno Almeida
NOVA	Internal Reviewer	Pedro Maló
AGRO	Internal Reviewer	Gregory Mygdakos
INNOV	Quality Assurance	Dimitris Drakoulis

Revision History

Version	Date	Partner(s)	Description
0.1	2021-02-18	UPRC	ToC & baseline
0.2	2021-03-02	UPRC	Description of back-end functionality and specifications
0.3	2021-03-17	INNOV	Contributions mainly to Section 4.2
0.4	2021-03-22	UNP	Contributions mainly regarding the presentation layer
0.5	2021-03-24	UPRC	Updates in several sections
0.6	2021-03-24	CP	Contributions mainly in Sections 2.1, 2.3
1.0	2021-03-29	UPRC	First Version for Internal Review
1.1	2021-03-30	AGRO, NOVA	Review Updates
2.0, 2.1	2021-03-31	UPRC, INNOV	Version for Quality Assurance; Review by the Quality Manager
3.0	2021-03-31	UPRC	Version for Submission

[1] Lead Beneficiary, Contributor, Internal Reviewer, Quality Assurance

[2] Can be left void

Executive Summary

Several data scientists and data practitioners develop innovative algorithms to address the needs of various applications in different domains. The same applies in the finance / insurance domains, with machine and deep learning algorithms being implemented and released by different experts. In INFINITECH project, a number of such experts are part of the consortium and will implement, validate and release the respective algorithms for the variety of the project's pilots. In this context, one of the key exploitation paths of the project is the release of these algorithms, as well as of bundle of algorithms acting as data analytics pipelines (i.e., ready-to-use solutions) to the wider research and innovation community. To realize the latter, this deliverable introduces the INFINITECH market platform, which will hold and offer the aforementioned machine and deep learning algorithms. As a marketplace, the current design provides information on how several actors (and contributors) can interact with the marketplace in order to ingest and retrieve the respective algorithms. Moreover, it introduces the concept of assets, since the INFINITECH marketplace will go beyond algorithms and will store and offer additional elements, including rich descriptions of the algorithms, validation datasets, training and evaluation outcomes for these algorithms. The market platform will also act as a digital innovation hub by hosting innovation management services, the so-called Virtualized Digital Innovation Hub - VDIH services that can be exploited by FinTech and InsuranceTech practitioners. These services reflect an additional proposition (as offerings) of the market platform that complements and provides additional value to the assets described above (e.g., algorithms, validation datasets, etc).

The INFINITECH marketplace specification exemplifies the respective layers of the marketplace, the functionalities supported by each layer, the interaction points with different stakeholders as well as the technical details that will drive the implementation of the marketplace. Furthermore, it reviews the current state of the art in order to identify the baseline technologies and approaches for the realization of the marketplace.

The current report is a follow-up on the specifications of the INFINITECH market platform as defined in the previous version of the deliverable (i.e., D8.1). Several updates have been introduced in the deliverable, including among others the completely new design specifications of the presentation layer, the assets to be offered by the market platform, additional use cases to be supported. Based on these updated specifications, the INFINITECH market platform is realized and the initial version is released ahead of schedule (based on the DoA planning) with a number of assets made available through the market platform.

Table of Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
BDV	Big Data Value
BDVA	Big Data Value Association
CRUD	Create Retrieve Update Delete - Basic Operations in DBMS
DIH	Digital Innovation Hub
DIS	Digital Innovation Services
DL	Deep Learning
EU	European Union
HTTP	Hypertext Transfer Protocol
IoT	Internet of Things
IPRs	Intellectual Property Rights
IRA	INFINITECH Reference Architecture
ISO	International Organization for Standardization
JSON	JavaScript Object Notation
ML	Machine Learning
R&D	Research & Development
RA	Reference Architecture
REST	Representational State Transfer software architectural style
SES	Sandboxes Experimenting Services
SME	Small and Medium-Sized Enterprises
TRL	Technology Readiness Levels
UI	User Interface
VDIH	Virtualized Digital Innovation Hub

1. Introduction

INFINITECH will provide a complete integrated environment enabling the utilization of big data and AI techniques in the finance and insurance sectors. The latter will be feasible through a set of technologies that enable exploitation of various datasets (obtained from different sources), optimized data management for these datasets (e.g., across diverse data stores), analytics with innovative algorithms covering a wide set of scenarios in the finance and insurance sectors, as well as use of tailored sandboxes on the underlying infrastructure layer for the execution of the aforementioned algorithms. These algorithms will leverage big data datasets that are available in the financial organizations of the consortium, high velocity data from IoT devices (i.e., connected cars, medical devices, smart phones), along with alternative data from a wide array of open sources like news and social media.

The INFINITECH solution goes beyond the utilization of analytics on specific datasets for a number of pilots / use cases, by aiming at a generalized approach that will facilitate the exploitation of various analytics algorithms (provided both by INFINITECH researchers / partners and by 3rd party data analysts) on top of different datasets. To this end, the analytics algorithms will need to be made available, to be described in terms of functionality, parameters and offerings, to be accompanied with datasets that can be used by interested parties in order to validate their applicability and performance and to be offered as ready-to-be-executed solutions (e.g., containerized) in order to increase their utilization. All these are representative functionalities of the INFINITECH market platform, which will hold and offer the solutions for realizing big data and AI techniques in the finance and insurance sectors.

Based on the above, the INFINITECH's multi-sided market platform aims at being one of the project's main ambassadors to the big data and AI communities. It will be a single, public and hybrid system with many different APIs, in order to cover all the different required perspectives of the platform. The market platform will offer big data and AI solutions, as well as IoT and Blockchain solutions, and VDIH Services. Thus, the INFINITECH market platform will be a four-perspective, unified environment being able to store several types of assets (e.g., algorithms, descriptions of algorithms, evaluation and validation results, datasets, experimentation outcomes, etc.) in any format. Therefore, it will be a single endpoint to various stakeholders, while also being open to the big data and AI communities to contribute algorithms and services.

1.1. Objective of the Deliverable

The main objective of this deliverable is to provide the updated architecture specifications of the INFINITECH market platform. The current deliverable is an updated version of the deliverable D8.1 (entitled "Market Platform and VDIH Specifications - I") and thus, it extends it with respect to the updated specifications based on the reflections of the initial design and implementation of the market platform. Based on these specifications, the realization of the market platform (i.e., its implementation) is realized (an initial version), as well as its population with the respective assets (e.g., algorithms, experimental results, etc.). The main updates of this deliverable compared to the previous one (D8.1) are described in section 1.3.

1.2. Structure

This document is structured as follows: The remaining of this chapter lists the differences of this report compared to D8.1 and introduces relevant approaches that serve as market platforms in order to identify potential links and baseline technologies for the INFINITECH marketplace. Section 2 describes the actors interacting with the marketplace and the offerings towards these actors, while Section 3 provides a high-level overview of the marketplace architecture correlating its main blocks with the INFINITECH and the BDV reference architectures. Section 4 is the core section of this report, presenting in detail the architecture of the market platform. The report concludes with a reference to the presented material in Section 5.

1.3. Updates with Respect to the Previous Version

This deliverable is an updated version of D8.1 entitled "Market Platform and VDIH Specifications - I". The main changes are reflected in the Presentation Layer of the market platform as it was decided not to use the IoT-Catalogue as a platform that will present the assets and the results of the INFINITECH marketplace, but with this as a basis, a new presentation layer will be developed, specially designed for the needs of INFINITECH project and the more specifically of the assets that will be hosted and offered by the market platform. Section 4.1.4 has been heavily updated to reflect the aforementioned change.

In addition, the list of assets that will be offered by the market platform has been updated (section 2.3.1 "Offerings"), following an analysis performed in the consortium regarding the assets that will be offered through the market platform and the respective interests expressed by the partners. Additional material regarding VDIH specifications has been added in section 4.2, as well as additional use cases that are described in section 4.3. Finally, the subsections describing the solutions per field / scope of use were moved from section 3.1 to section 2.1, and the section 3.1 has been extended regarding the structure of the marketplace and its main components.

1.4. Relevant Work in Marketplaces and Potential Synergies

This section summarizes relevant works that have been identified and can be exploited as a baseline to realize the INFINITECH's market platform, with additional information that is included in the previous version of the market platform specifications (i.e., D8.1).

1.4.1. The FINSEC Marketplace

The H2020 FINSEC project has established and launched the Finsecurity.eu (https://finsecurity.eu) market platform. It serves as a promotional channel for the solutions and services that have been developed in the FINSEC project, while providing access to relevant knowledge assets such as whitepapers and training presentations about security and digital finance services. The assets that are included and promoted in the platform are clustered in four main categories: (i) solutions, (ii) services, (iii) demonstrators and case studies, (iv) digital finance academy for security, and (v) blog & news. Finsecurity offers selected contents to registered members only.

Relevance and Synergies with the INFINITECH Market Platform:

There is a direct and close thematic relevance between Finsecurity and the INFINITECH market platform. Hence, the following synergies will be explored:

- **Content sharing between the two platforms**, as items in one platform are relevant to the community of the other and vice versa. Hence, content such as solution descriptions, knowledge assets and presentations could be shared between the two platforms (e.g., linked from one platform to another and vice versa).

- **Joint community building and cross-registration**, as both platforms are addressed to users of the digital finance community. Some members of Finsecurity are likely to be willing to join the INFINITECH market platform and vice versa. This could boost the community building efforts for both projects (i.e., FINSEC and INFINITECH). In this direction, the two projects will also explore mechanisms for incentivizing participation in both platforms e.g., by providing premium content that will be accessible to registered participants to both platforms only.

1.4.2. The KNIME Hub

The KNIME Hub (https://hub.knime.com) is a one-stop shop where data scientists and machine learning experts can find and collaborate on workflows and nodes that work over the popular KNIME.com open analytics platform. It provides its users with solutions to their data science questions. The hub provides a powerful search engine that enables users to find data sciences resources in the form of modular components and relevant solutions in the form of end-to-end pipelines that combine multiple modules / components.

Relevance and Synergies with the INFINITECH Market Platform:

INFINITECH is a big data and Data Science project, destined to produce data science components such as components for data processing, data analytics and machine learning. Therefore, the KNIME Hub concept for accessing different modules, viewing information and examples about them, while ultimately using and extending them in the KNIME platform, is very relevant to INFINITECH. The offering of INFINITECH market platform shall include INFINITECH modules in a similar fashion, including modules that can be downloaded and run over the INFINITECH big data platform or any other data science platform that will support INFINITECH.

1.4.3.   The IoT-Catalogue

The IoT Catalogue (https://www.iot-catalogue.com) is an already available web-based catalogue and decision-support tool for solutions of the Internet-of-Things (IoT). The 'IoT Catalogue' targets specifically developers / integrators of IoT systems addressing questions such as: What IoT solutions exist for a given problem? What components compose a given IoT solution? What is their cost? Where to buy them from? etc. All the components used in a solution are represented with detailed information such as manufacturer, product page and its vendors and allow the user to choose where to buy based on the store location, price, etc. The components are categorised in different types being type-specific information added to each component. In this tool, several solutions can be considered when taking into account different environments and their specific requirements.

Relevance and Synergies with the INFINITECH Market Platform:

The INFINITECH market platform has as a reference the IoT-Catalogue to provide to stakeholders, information related to the INFINITECH's related technologies and also the internal Use Cases coming from the INFINITECH's consortium. Moreover, the IoT-Catalogue provides a set of functionalities that can be explored by INFINITECH, such as:

- **Integration on 3rd party websites:** The IoT-Catalogue has the ability to be embedded in external pages. This function is accomplished using an IFrame. An IFrame (Inline Frame) is an HTML document embedded inside another HTML document on a website. This is often used to insert content from another source, it can be the whole webpage or certain parts as it is specified.

- **Development and Integration of 3rd party plugins**: Third-parties' sources of information can integrate with IoT-Catalogue through the REST API. The access to this API is protected by an API key for authentication and managing the authorizations to the access to the provided by the API. The IoT-Catalogue platform is composed of a collection of several customizable modules. These modules can then be customized by third-party entities to integrate new kinds of data and to extend visualization capabilities to show the corresponding information. As an example, in H2020 CROSSMINER, the module of a new webpage was defined to represent the information about the activity of open-source repositories and corresponding bug tracking system.

1.4.4.   The 5GTANGO Catalogue

5GTANGO (https://www.5gtango.eu) was an EU project whose purpose is 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. During the course of the project, a market platform - named as "Catalogue" has been developed with the main purpose to host and offer the relevant network-related services. The Catalogue is an instrumental component of the 5GTANGO environment, presenting different parts of it. Primarily, it hosts the different descriptors of the 5GTANGO packages. Since 5GTANGO aims at a multi-platform environment, it enables the developers to orchestrate and deploy their services using different Service Platforms. In this context, the 5GTANGO Catalogue is being adapted to support the storing and retrieval of new packages.

Relevance and Synergies with the INFINITECH Market Platform:

INFINITECH can exploit the functionalities of the catalogue in order to support the basic usage scenarios that are described in Section 2 of this deliverable. The main points / functionalities are the following:

- Support of **different types of assets on the storage layer**: The 5GTANGO catalogue employs an underlying storage layer that allows the storage of objects, files (on the file system) and structural documents (through a MongoDB database). Given that in INFINITECH various assets will also be stored in the same manner (e.g., algorithms, evaluation results, descriptors, etc), the 5GTANGO storage engine fits this purpose.
- Support for **descriptors of the stored assets**: The 5GTANGO catalogue enables the description of Virtualized Network Functions and Network Services (as chains of functions). These are rich JSON descriptors that include several properties. In the context of INFINITECH, these descriptors can be adopted and adapted to describe the stored assets.
- Support for **advanced search and retrieval of assets**: The 5GTANGO catalogue incorporates a set of mechanisms to support the search and retrieval of the stored entities based on their descriptors in a tree-structure, thus enabling the assets to be retrieved based on a number of properties. The latter is quite relevant for INFINITECH since search for algorithms should be supported for all their properties (e.g., category, input data formats, etc).
- Support for **various interaction methods**: The 5GTANGO catalogue includes a set of APIs to enable both the ingestion and the retrieval of assets from different components. These APIs can be exploited in INFINITECH and extended to meet the stakeholders' needs in terms of ingestion of algorithms (e.g., from data scientists) or browsing and retrieval of algorithms. Given that a richer interface is required, the 5GTANGO Catalogue can be combined with the IoT-Catalogue with an intermediate API that will facilitate communication and thus promote interoperability.

1.4.5. Selection of baseline

Based on the aforementioned descriptions and analysis, for the INFINITECH market platform, in the scope of the back-end solution (as described in Section 3), the 5GTANGO Catalogue has been utilized since it enables both persistent storage and descriptions of the assets to be stored, while also providing specific APIs for the relevant operations for Creating, Retrieving, Updating and Deleting (CRUD) descriptors. Additionally, it facilitates full-text search capability in structure-agnostic documents since the schema of the diverse documents (i.e., descriptors of different types of assets) is variable. Regarding the INFINITECH market platform front-end solution, the knowledge from the IoT-Catalogue in terms of building the interface as well as enabling information and data retrieval through respective APIs, has been utilized and extended given the different types of assets that will be offered through the INFINITECH market platform.

2. Consumers and Offerings

This section has been updated to summarize the main actors that interact with the INFINITECH marketplace along with the respective user journeys. Moreover, it enlists the different categories of solutions that will be available through the platform and the types of offerings.

2.1. Solutions per field of use

The following subsections summarize the solutions that will be available through the INFINITECH market platform.

2.1.1. Big Data and AI Solutions

Big data is a key area for businesses and the public sector analysing huge amounts of data to discover new, innovative ideas, technologies and solutions. By utilizing big data analytics and artificial intelligence, businesses and organizations can support their decision making (i.e., through data-driven insights) depending on the domain being addressed. Additionally, they can exploit big data and AI solutions as offerings towards their customers either explicitly (e.g., IT companies towards other industries such as retail) or implicitly (e.g., by analysing the customers' behaviour to provide personalized services and products).

Several categories of solutions, i.e., offerings, correspond to the assets described in Section 2.2 of this document. For example, algorithms and datasets are essentially the basis of any big data and AI solution, where an algorithm can use and process a multitude of data in order to produce results in the given application context, such as detecting fraud. An extension of these solutions are cloud containers (e.g., Docker containers), which are isolated environments where someone can package and run applications directly within the host machine's kernel. These containers are an excellent choice for continuous integration and continuous delivery workflows.

In INFINITECH project, a number of big data and AI solutions will be researched and applied to the respective pilots of the project. As described in deliverable D2.5 (entitled "Specifications of INFINITECH Technologies - I") these solutions reflect analytics and ML/DL algorithms that will be stored and offered through the INFINITECH marketplace. The solutions have been analysed in terms of their structural elements, such as their inputs and outputs, what technologies will be used for these parts, licenses, etc. Based on this analysis, the solutions will form the assets of the INFINITECH marketplace, and their descriptions and analysis outcomes will be reflected in the structural representations (i.e., JSON descriptors of the assets) that are also stored in the INFINITECH marketplace. For example, the fraud detection algorithm / service will be stored as an asset, while both the inputs and the outputs of the service, as well as the trained model will be stored. The same applies for example for the savings proposal recommender, for which besides the recommender per se, sample data that will also be stored in the marketplace, which can be used by the consumers of the marketplace as execution examples and evaluation of the functionality of the recommender. It is expected, that during the course of the project all developed solutions will be part of the marketplace along with additional information as described in the examples above (e.g., input / output information in the descriptor, execution examples, etc).

2.1.2. IoT and Blockchain Solutions

The vast majority of digital transformation applications for the finance and insurance sectors are data intensive. This holds for applications in different areas such as retail banking, corporate banking, payments, investment banking, capital markets, insurance services, financial services security and more. All these applications leverage very large datasets from legacy banking systems (e.g., customer accounts, customer transactions, investment portfolio data), which they combine with other data sources such as financial markets data, regulatory datasets, real-time retail transactions and more. With the advent of Internet-of-Things (IoT) devices and applications (e.g., Fitbit, smart phones, smart home devices), several FinTech and InsuranceTech applications can take advantage of contextual data in order to offer better quality of service at a more competitive cost (e.g., personalized healthcare insurance based on medical devices and improved car insurance based on connected car sensors). Furthermore, alternative data sources (e.g., social media and

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

INFINITECH has a clear objective related to exploitation of IoT and Blockchain technologies in the finance and insurance sectors. There already exists a wide set of technologies researched and identified, that can become of use within the scope of IoT and Blockchain solutions. These lists are provided in INFINITECH's "D2.6 – Specifications of INFINITECH Technologies – II", where a section is dedicated to the Component Group: "Blockchain and Information sharing". In this section, there exist several blockchain related technologies that are provided by partners within INFINITECH's consortium and are of interest for the INFINITECH's pilots. During the course of the project, WP8 will work with the providers of these technologies and standardize the way that these technologies will be made available through INFINITECH marketplace (Tasks 8.2 – 8.6).

2.1.3. Virtualized Digital Innovation Services and Resources

The INFINITECH VDIH offers key accelerations services and resources for companies. The main objective of such services is to support businesses in their digital transformation leveraging on the expertise that the INFINITECH consortium can utilize and exploit. The VDIH is a Virtualized shop based upon technological infrastructure and brings knowledge and experience to provide services such as piloting, training and consulting to implement innovative solutions across the value chain.

The specific services will be defined and prioritized, by taking into account innovative business models to implement each of them, as well as adhering to IPRs and commercial agreements. The provided services will be standardized and federated after defining ontologies and setting blueprints.

Moreover, the service federation will follow a design thinking approach, mapping and integrating the other INFINITECH services and the other third-party services.

2.1.4. Third party solutions

The INFINITECH market platform will be a repository of all the project's results, which will be made available to the EU digital finance/insurance communities. The project will provide the means for integrating within the platform, not only solutions of the consortium partners, but also solution from third-party solution providers and innovators (e.g., solutions that will be implemented during the project's hackathons or solutions that will be developed as part of the INFINITECH pilots).

The DIH will serve as a single-entry point for accessing INFINITECH resources for innovation in the finance/insurance sectors (including FinTech/InsuranceTech). It will aggregate resources from FinTech accelerators/clusters, incubators of the incumbent financial organizations, FinTech developers, R&D organizations and other stakeholders of the consortium. It will leverage resources from the project's tailored sandboxes, including the EU-wide experimentation infrastructure that the project will establish for FinTech & InsuranceTech firms, which will be enhanced with complementary assets and services such as training, consulting and business support.

Various third-party solutions in the areas of Regulatory Compliance (e.g., Anonymization, Digital On-boarding, etc), Natural Language Processing, Sentiment Analysis, Insurance or Investment Products Recommendation Engines, etc) will be available through INFINITECH Project. Also, other solutions that will be developed from Fintech or InsuranceTech, and that will benefit from the outcomes of the INFINITECH Big Data, IoT and AI technologies, will also be included in order to provide alternative applications related to the Financial or Insurance Industry. The list of the available third-party solutions that will be available through INFINITECH market platform, will be placed in a separate deliverable (D8.9), grouped by the relevant functionality and sector that will be mainly relevant to.

For all third-party solutions, that will be part of the marketplace -along with additional information for each one- specific examples will be included, as well as the relative Open API input / output information descriptors and relative sandbox and datasets that may be available either through INFINITECH, or other sources (e.g., Open Data, Synthetic Data, etc).

2.2. User Journeys

The following tables illustrate potential user / customer journeys. These journeys could facilitate understanding of what kind of services and functionalities should be offered by the platform. The different users that have been selected include a diversity in terms of roles. They cover: (i) a FinTech developer as a role that aims at exploiting the assets available in the INFINITECH marketplace in the scope of evaluating different solutions (e.g. analytics algorithms or workflows of algorithms) on various datasets, thus mainly being a consumer of the offerings with an emphasis on utilizing not only algorithms but also datasets in the overall evaluation process, (ii) a digital finance consultant as a role that aims at using the assets available in the INFINITECH marketplace and applying them in a real-world case, thus going beyond the experimentation and validation to actual use, and (iii) an INFINITECH developer as a role that represents the providers of the solutions (through the marketplace) to showcase how solutions can be ingested and be made available.

Table 1 - A FinTech Developer Customer Journey.

Persona Definition
FinTech Developer ▪ Peter is a FinTech developer that works on a FinTech company that develops a novel wealth management product. ▪ He has a BSc in Software engineering and an MSc in Finance. ▪ As part of his start-up development activities, he is interested in novel data-driven wealth management solutions. ▪ His company is seeking to incorporate machine learning and big data management in its wealth management product.
Goals & Objectives of the Persona
Objectives Peter wants to gain knowledge and skills that will ensure his successful engagement in the wealth management developments above. He is also interested in accessing data science examples that could inspire his developments in the company. Therefore, he would like to: ▪ Understand the basics of data science for wealth management in digital finance. ▪ Gain insights on the use of machine learning pipelines and data science workflows for wealth management in his product / project. ▪ See some practical use cases of such solutions i.e., financial organizations improving their wealth management approach based on Machine Learning solutions over large wealth management datasets.
Touchpoints in the INFINITECH Marketplace
Touchpoints and Dissemination Channels Peter is aware of the INFINITECH marketplace as a single-entry point for accessing big data and AI resources for digital finance in Europe. This awareness is a result of one or more of the following: 1 Peter was informed about INFINITECH marketplace during a conference, as part of his discussions with a colleague. 2 Peter read information about INFINITECH in a widely spread Social post, written by an influencer of the FinTech community.

3	Peter attended one of the hackathons of the INFINITECH consortium, organized by Copenhagen FinTech.

INFINITECH Functionalities in the Persona's Journey
Journey ▪ Peter visits the home page of the INFINITECH platform. ▪ He reads about INFINTECH and then selects to access the "Digital Finance Academy" and "Workflows" menus. ▪ Under the "Training Presentations" menu, Peter accesses a 10' tutorial webinar about Wealth Management in Digital Finance using Machine Learning and big data Management Techniques. He has registered with the INFINITECH Marketplace platform in order to access the webinar and to download relevant materials. ▪ Peter can also read a post with a simple example on how to build a digital finance pipeline in the INFINITECH platform. ▪ Peter visits the INFINITECH marketplace few days later. This time he accesses the "Data Workflows" section, where he searches about wealth management data science components and solutions. There he can find relevant INFINITECH workflows in the form of complete, end-to-end pipelines consisting of INFINITECH components. ▪ Peter is able to see the documentation of individual components as well as of the integrated pipeline. He can also read about how the pipeline is used in INFINITECH pilots. Finally, he can execute simple demos of individual modules or even the end-to-end pipeline. The demos are based on sample datasets (e.g., customer data, investment data) that are anonymized and available through the market platform.

Table 2 - Customer Journey for a Digital Finance Consultant.

Persona Definition
Digital Finance Consultant ▪ Marta is a Digital Finance Consultant for Banks. ▪ Marta needs cutting edge technology and wide scope of solutions to propose the more appropriate and suitable to her client's needs. ▪ Marta has recently received requests from her clients, for intelligent finance management solutions for retail customers. In collaboration with her colleagues, she has started reviewing some of the promising and challenging approaches for data-driven financial management for individuals and businesses (notably SMEs).
Goals & Objectives of the Persona
Objectives Marta wants to gain access to existing solutions for intelligent finance management for SMEs, notably solutions that will help SMEs plan their cash flows and ensure their liquidity. Marta wants to evaluate such solutions and to propose them to her customers (if applicable). Specifically, she is interested in: ▪ Finding detailed descriptions and the benefits of each solution. ▪ Exploring technical details about the solutions including information on data needed, machine learning models and more.

▪ Finding out practical use cases of such solution in relevant organizations, such as banks.
▪ Being able to see some demonstrable and working software solutions.
▪ Being able to contact the vendor of the solution.

Touchpoints in the INFINITECH Marketplace
Touchpoints and Dissemination Channels
Marta is aware of the INFINITECH marketplace through one or more of the following channels:
▪ Marta saw a presentation about INFINITECH in a digital finance conference.
▪ Marta's colleagues work with financial organizations, which are already using some of the INFINITECH solutions and services.

INFINITECH Functionalities in the Persona's Journey
Journey
▪ Marta visits the INFINITECH marketplace and notices the ability to search modules, workflows and entire data-driven solutions.
▪ Marta searches relevant solutions and offerings based on keywords like "Personal Finance Management" and SMBs.
▪ Marta finds relevant INFINITECH components and pipelines, including demos and sample workflows of the INFINITECH solutions that were developed for the BOC pilot.
▪ Marta is able to access detailed documentation about these solutions and about the individual modules that they comprise.
▪ In the "Forum" and "Reviews" sections, Marta can find more information and join the discussion of the community on relevant topics.
▪ Marta is thinking of downloading the workflow for a fee, based on a PayPal payment. She also sees who she can contact for more information about the selected solutions. She will be considering partnering with INFINITECH partners towards offering similar solutions to her customers.

Table 3 - An INFINITECH Developer Journey.

Persona Definition
INFINITECH Developer
▪ Michael is an INFINITECH developer who is responsible for the development and maintenance of one of the project's pilots.
▪ During his work on the pilot's resources, he discovers an alternative approach for an offered solution of the pilot.
▪ He and his associates decide that the new approach should not replace the old one, and thus they make the new approach available as an extension.
Goals & Objectives of the Persona
Objectives
Michael wants to upload his "discovery" on a platform from where the users of the pilots will be able to easily retrieve it and then use it in the pilot. He would also like to:
▪ Limit the access to his solution and make it available only to people who use the pilot for which he works.

▪ Inform, the end users properly about what needs are covered by the new approach, how it should be used, what are the expected results of its execution and what are the inputs. ▪ Be able to receive feedback from the users about their experience with the new solution.
Touchpoints in the INFINITECH Marketplace
Touchpoints and Dissemination Channels
Michael knows that INFINITECH for which he works, has its own marketplace. Throughout the duration of the project, he is informed about its progress and the functionalities it offers, through the presentations in conferences, the relevant deliverables that concern it (such as the present document) but also by internal presentations.
INFINITECH Functionalities in the Persona's Journey
Journey ▪ Michael knows that the INFINITECH market platform can cover his needs. ▪ He visits the home page of the INFINITECH platform. ▪ He logs-in with his credentials to the platform (or creates an account if he has not already one). ▪ He chooses to upload a new asset and he is being redirected to a page where he will have to fill in some necessary information about his asset. ▪ On the same page, he will be able to limit the access to his file but also to add information that will inform the end users for its usage and more. ▪ After the completion of all required fields, he submits his asset and if everything is correct, he can disconnect from the platform. ▪ A few days later, some end users who executed his solution on the pilot, reverted to him with some feedback. ▪ Michael is being informed of their feedback through his email and logs-in to the platform in order to see or respond to them.

2.3. Offerings, Providers and End-Users

This section maps the end-users of INFINITECH's market platform and provides an overview of the platform's offerings to them, reflecting also potential providers.

2.3.1. Offerings

With the term offerings, we describe all the assets (objects) that the INFINITECH market platform will support and therefore every asset that will be stored on it and will eventually be able to be retrieved by the end-users. The offered assets of the market platform vary in their field of use. There will be assets that will cover and provide solutions to the especially popular fields of Big Data and Blockchain, to the world-wide known field of AI and to the rapidly increasing field of IoT. Examples of the market platform's stored assets, are the following:

- Algorithms, which can be executed after being retrieved (e.g., ML / DL algorithms). It should be noted that these actually refer to a "single" component, not to a composite one / workflow that is addressed in the next point below.

- Analytics pipelines, which refer to a set of algorithms providing a specific solution (e.g., fraud detection). Pipelines can also include data preparatory mechanisms (e.g., for data cleaning) along with the algorithms / analytics that have been described above. It should be noted that these actually refer to a "full solution" / "composite" component search, retrieval and utilization.

- ▪ <u>Datasets</u>, which are utilized by algorithm in order to produce interesting results. Moreover, the datasets can be used as training datasets to ML/DL algorithms, etc.
- ▪ <u>Machine Learning models and demos</u>, which refer to the model artifact that is created by the training process of a machine learning algorithm.
- ▪ <u>Python Jupyter notebooks</u>, which are interactive computational environments, in which users are able to combine code execution, rich text and media.
- ▪ <u>Scientific studies and / or tutorials (training courses or training presentations)</u>, which could train the end-users and to help them to achieve their objectives.
- ▪ <u>Webinars and / or lectures (videos)</u>, which could be outcomes of INFINITECH's events or even provided by the INFINITECH partners (e.g., universities could provide some videos from their courses).
- ▪ <u>Experimentation results</u>, which refer to the outcomes of executions of the previous algorithms on specific datasets, or outcomes of new studies, etc.
- ▪ <u>VDIH services</u>, which they are going to enrich Europe's DIHs by enhancing the project's VDIH. The DIHs are one-stop-shops that support businesses in their digital transformation.
- ▪ <u>Docker containers</u>, which are pre-build, standalone and executable packages of the algorithms or pipelines (i.e., set of algorithms) that can be directly exploited by the relevant stakeholders.
- ▪ <u>Combined solutions</u>, which in fact are combinations of the above types of assets, considering them as groups of assets towards the provision of solutions for different types of stakeholders and actors as described in section 2.3.3.

All these assets emerge from the needs identified by different INFINITECH consortium partners (technical and pilot) and have been extended with extra suggested assets as a fruitful way to exploit and disseminate the INFINITECH results to the community. More specifically, the list above has been formed based on various inputs that have been collected in the scope of WP8 but also of WP2, and specifically documented in D2.6 (entitled "Specifications of INFINITECH Technologies – II") that analyses INFINITECH components and tools.

2.3.2. Providers

Given the variety of offerings / assets in the INFINITECH marketplace, different providers are envisioned. Providers can upload and at the same time publish to INFINITECH's market platform their assets. If the providers are the authors of the provided asset, they can also be called producers. The following are the basic types of providers:

- ▪ <u>Owners</u>, as actual developers who created an asset (e.g., a data scientist of the project can upload his / her algorithm or a complete solution – set of algorithms). Owners can also be scientists as described below but the key differentiating actor is that different business objectives (e.g., in terms of sharing) may also be specified by owners.
- ▪ <u>Scientists</u>, who contributes assets that would like to share with other scientists for the same reason as in the case of owners. Besides algorithms or pipelines, experimentation results are some additional examples of assets that can be uploaded to the market platform by a scientist.
- ▪ <u>Services</u>, where new assets, for example datasets, are generated and stored on the market platform during or after their execution. An example of a service – provider could be a framework that collects results from ML algorithms and stores them on the platform.

More examples to the above types of providers are the INFINITECH's members e.g., technical members, analysts, etc. that fall in the classes of owners or scientists as described above, other authorized users, e.g., members from other collaborating projects who might be engaged with INFINITECH due to ISA (INFINITECH Stakeholders Alliance).

These can contribute and enhance the offerings of the market platform. Especially for the third parties, their solutions that will be uploaded to the market platform can be the outcomes of hackathons, innovation contests and other events that will be organized by the project's consortium.

Finally, core providers to the market platform are the VDIH services providers, who will enrich the VDIH perspective, with new solutions.

2.3.3. End-Users

The end-users of the market platform, also named as consumers, are mainly the FinTech and InsuranceTech communities. These communities will aim at performing analytics on their environments – in the context of the project these environments are the INFINITECH's sandboxes and testbeds. To this end, the end-users will be able to retrieve the offered assets and / or services for which they are interested and use them by deploying and executing of them, depending on their needs. Thus, the end-users might be: (i) data scientists that would like to obtain machine learning and AI algorithms, experiment and evaluate them (utilizing also the provided datasets and experimentation results offered by the marketplace), (ii) solution providers that retrieve assets and apply them (either as they are or through additional enhancements on them) directly to specific cases in the insurance and financial sectors, (iii) infrastructure providers that retrieve assets, deploy them on the infrastructure resources (especially in the case that containers are obtained) and deliver them as running, ready-to-use solutions.

The INFINITECH services i.e., the sandboxes that will host services from the different pilots of the project, can also be consumers of the described offerings and assets of the marketplace, by directly obtaining the respective assets. It is expected that they will obtain them through APIs in the form of containers, so that they can be directly deployed and utilized.

Moreover, since the services that will be developed for the VDIH that will be stored on the platform, the VDIH itself can be considered as a service – consumer because it will have to retrieve the services from the platform. Other platforms / services (e.g., "ALIDA micro-service platform"), will be able to retrieve services like docker containers and analytics pipelines through a bridged API.

Finally, the market platform is intended to serve the needs of the wider public – the big data and AI research and innovation community, so as to allow everyone to join the market platform and obtain the respective assets.

3. Architecture Overview

This section provides the overview of the market platform architecture, including its capabilities, functionalities and how the overall architecture links with the BDVA reference architecture.

3.1. Main Components and Structure

The INFINITECH market platform provides a number of functionalities which are mapped to different layers (as depicted in Figure 1 and Figure 5) that realize the expected capabilities as presented below, and further discussed in greater detail in Section 4:

- The "Interaction Layer", which supports the communication between the market platform and the end-users / consumers of the INFINITECH's ecosystem.
- The "Presentation Layer", which provides the User Interface (UI) towards different types of stakeholders.
- The "Assets Management Layer", which is responsible to deliver principles and techniques for the management of the market platform's assets / offerings.
- The "Assets Storage Layer", in which the platform's offerings will be stored.

Moreover, the market platform is structured around two core services, the back-end and the front-end. This approach contributes towards the platform's enhancements in terms of functionality as well as provides additional information and capabilities (e.g. in the case of a 3rd party willing to access the stored assets through the back-end). As described in section 2.3.3, the marketplace supports both end-users and other services (through the respective interfaces). In this context, the end-users are able to interact with the market platform through the front-end (through the presentation layer) that reflects a user-friendly platform (providing the UI), while other additional services (e.g. from 3rd parties) may interact directly with the back-end (through the interaction layer). This separation is also represented in Figure 12.
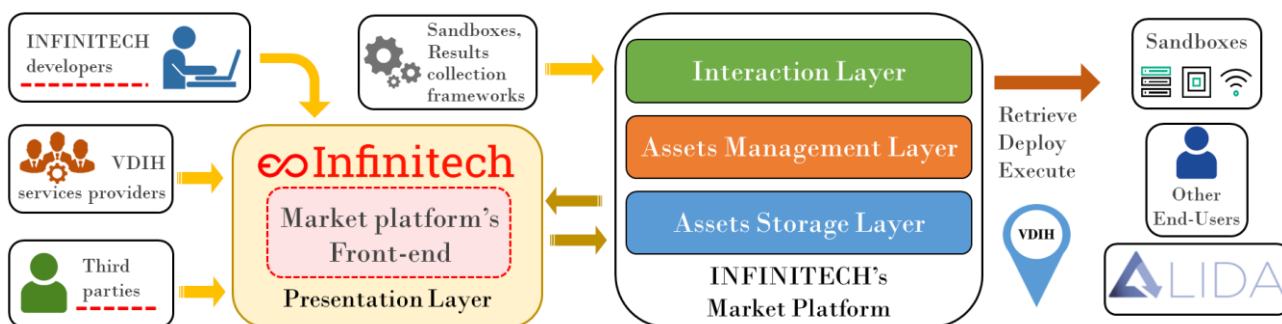

Figure 1 - INFINITECH's Market Platform including the Front-end and the VDIH.

As described in section 4.1, the core base of the marketplace (the back-end) is developed (an initial version has already been released), using a variety of components / tools. All of its sub-components are containerized as Docker images (https://www.docker.com) that among others offer more efficient management and maintenance (enabling continuous updates and integration). Python 3.8 (https://www.python.org) is used as the programming language that along with the Flask (https://flask.palletsprojects.com/en/1.1.x) frame-work, which is a Web Server Gateway Interface (WSGI) developed in Python, implements RESTful APIs to handle the respective HTTP requests.

The assets (offerings) are stored in a MongoDB database (https://www.mongodb.com) that is used in combination with GridFS specification (https://docs.mongodb.com/manual/core/gridfs) - for storing and retrieving large files / objects, of any format. Moreover, Gunicorn (Green Unicorn - https://gunicorn.org), a Python WSGI HTTP Server for UNIX is utilized with NGINX (https://www.nginx.com), an open-source high-performance HTTP web server and reverse proxy, since Flask is not optimum for production mode, and thus,

both tools extend the Flask framework in order to enable access to multiple users at the same time. Finally, the front-end exploits WordPress (https://wordpress.com) for the realization of the UI.

3.2. Link with BDVA RA

The BDV Reference Model has been developed by technical experts who are members of the Big Data Value Association (BDVA). This model may serve as a common reference framework to locate Big Data technologies on the overall IT stack and it addresses the main concerns and aspects to be considered for Big Data Value systems.

It is structured into horizontal and vertical concerns. We can imagine these concerns as layers but actually the BDVA is not considered as layered. The horizontal concerns cover specific aspects along the data processing chain, starting with data collection and ingestion, and extending to data visualization. The vertical concerns address cross-cutting issues, which may affect all the horizontal concerns such as Cyber-Security and Trust concerns. In addition, vertical concerns may also involve non-technical aspects.

The following figure (Figure 2) presents both the horizontal and the vertical concerns of the BDV's Reference Model which are described in BDVA's publication "BDVA Strategic Research and Innovation Agenda v4 (BDVA SRIA v4)" (http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf).
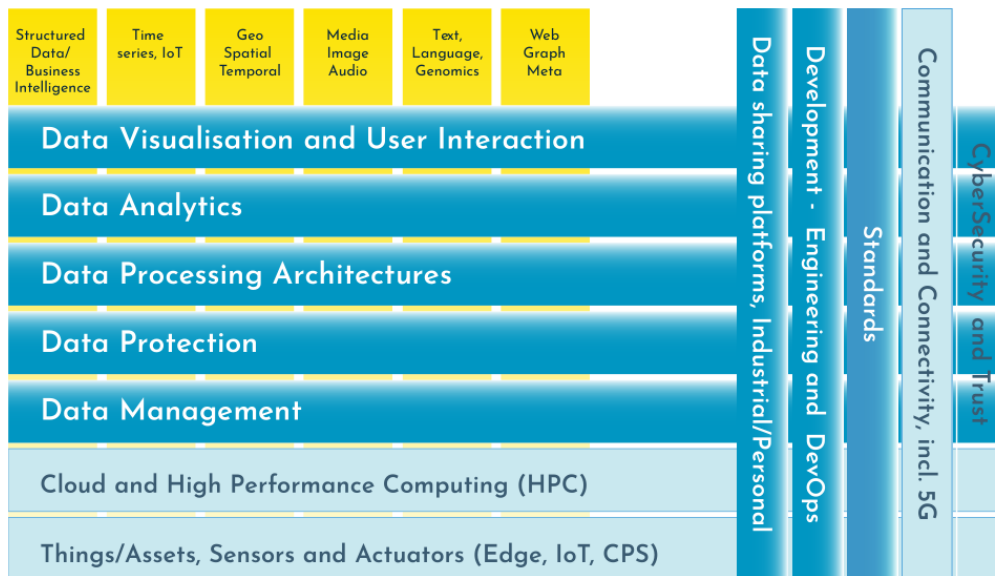

Figure 2 - BDV Reference Model.

The INFINITECH architecture is aligned and is compatible with the BDV Reference Model, as described in the respective INFINITECH Architecture deliverable (D2.13 - Reference Architecture – I) and shown in the following figure (Figure 3).
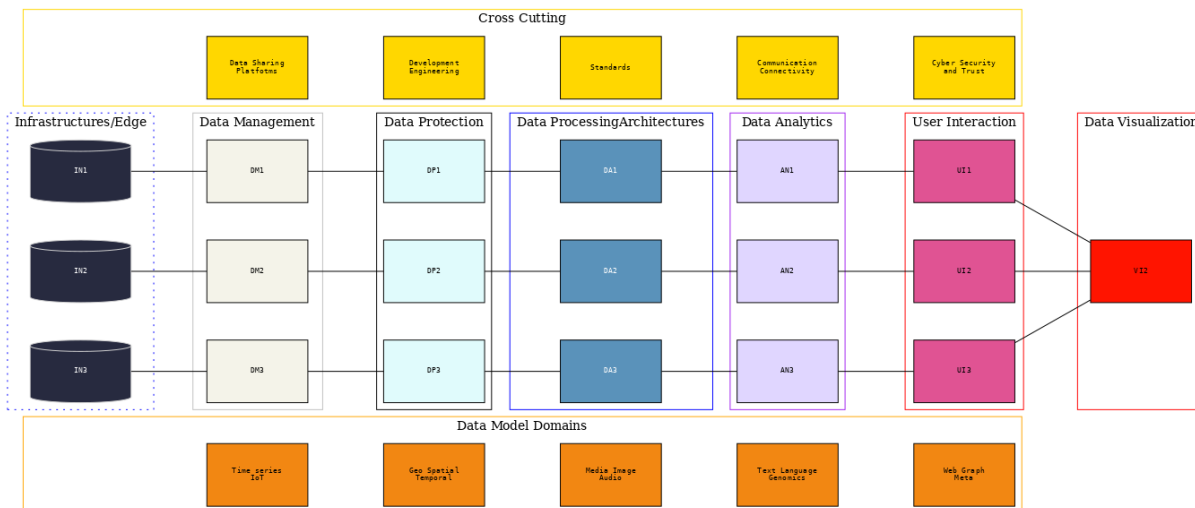
Figure 3 - INFINITECH Reference Architecture Mapping (IRA) with BDVA Reference Model.

Towards an integration of not only the overall INFINITECH architecture, but also of the INFINITECH marketplace (in order to foster its exploitation possibilities), the latter has been designed in such a way to be aligned with the BDV reference model as depicted in the following figure (Figure 4).
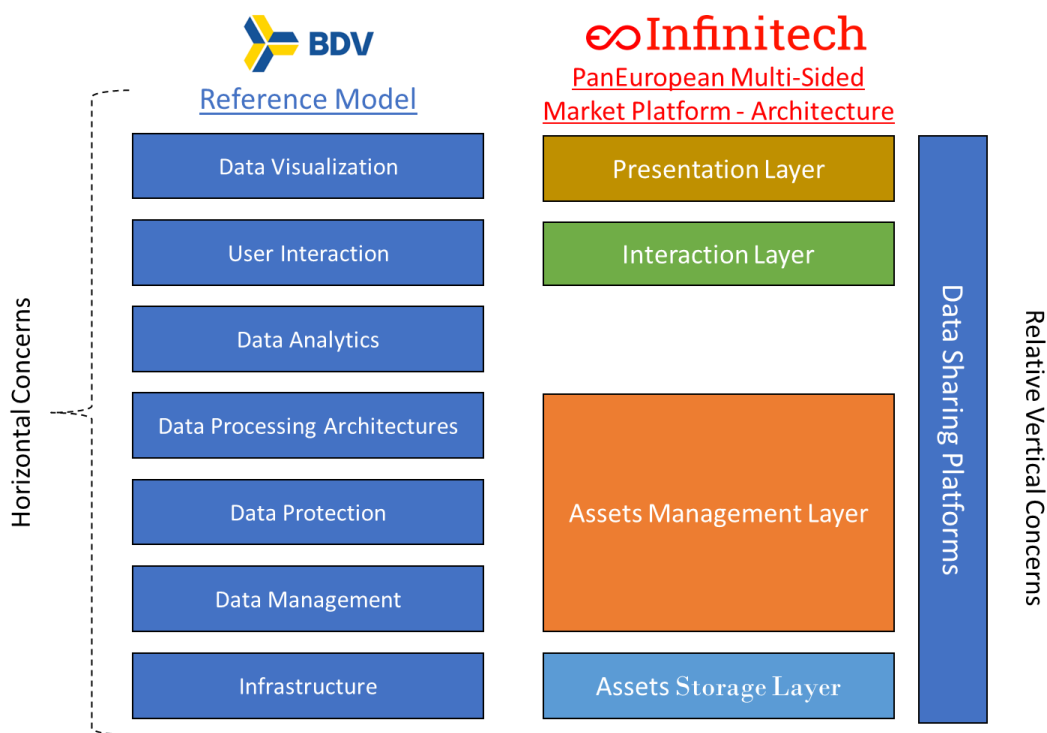

Figure 4 - INFINITECH market platform layers aligned with BDVA.

As shown in the figure, the main layers of the marketplace are aligned with BDV reference architecture as follows:

- The "Presentation Layer", which provides the user interface (UI), is actually related to the BDV's horizontal concern "Data Visualisation and User Interaction".

- The "Interaction Layer" that supports the communication between the market platform and the end-users / consumers of the INFINITECH's ecosystem (as described in Section 2) is mapped to the respective user interaction of BDV reference architecture.
- The "Assets Management Layer" of the INFINITECH marketplace is responsible for providing principles and techniques for the management of the stored assets. This layer maps to the data management of BDV reference architecture. As described in Section 4, the INFINITECH market platform will also encompass mechanisms for security and data protection, while also delivering the assets in a format that can be directly utilized for deployment and execution (i.e., as dockers). Thus, it maps to the respective layers of BDV reference architecture.
- The "Assets Storage Layer" is related to the vertical concern "Data Sharing Platforms" which includes the industrial and the personal data platforms along with the marketplaces. This means that the whole market platform can be considered as a vertical concern / layer of the BDV Reference Model.

Based on the above, as the entire INFINITECH Reference Architecture fully complies with the BDV Reference Model, so does and the INFINITECH Market Platform, with its layers referenced to the BDVA's layers.

4. Multi-sided Market Platform and VDIH Specifications

4.1. Multi-sided Market Platform Specifications

As outlined in the previous chapters of this deliverable, the INFINITECH multi-sided market platform will be a unified, public and hybrid system / environment with many different APIs. The platform will integrate its four-perspectives (hybrid) into a single system in order to be a more user-friendly platform, to reduce maintenance costs and to facilitate its management. Moreover, the market platform will be public to various stakeholders and to the big data and AI communities so contribution will be encouraged for the development of the services offered by the platform but also in the development of the communities to which it is addressed. Thus, the platform will also facilitate third party access. In terms of storage, the market platform will store various types of assets (objects) in any format, as described in Section 2.3.1 (i.e., Offerings), and furthermore relevant functionalities will be developed for the assets' management. As depicted in the following figure (Figure 5), the market platform includes four different layers, each one delivering specific functionalities, which will be further described in the next subsections.

In addition to these, the platform will take into account issues related to security for the assets it stores and offers. It will also consider, comply and inform the end users on privacy issues (i.e., GDPR). Hence, user registration is a required functionality so as to ensure that all the regulations are enforced. Thus, the providers of the assets, among other things, will be able to set rules on which end users will be able to retrieve their solutions, etc. For example, an INFINITECH developer, uploads to the market platform an algorithm intended for the INFINITECH sandboxes / pilots, and hence, sets the rule that the asset is only available for the sandboxes.
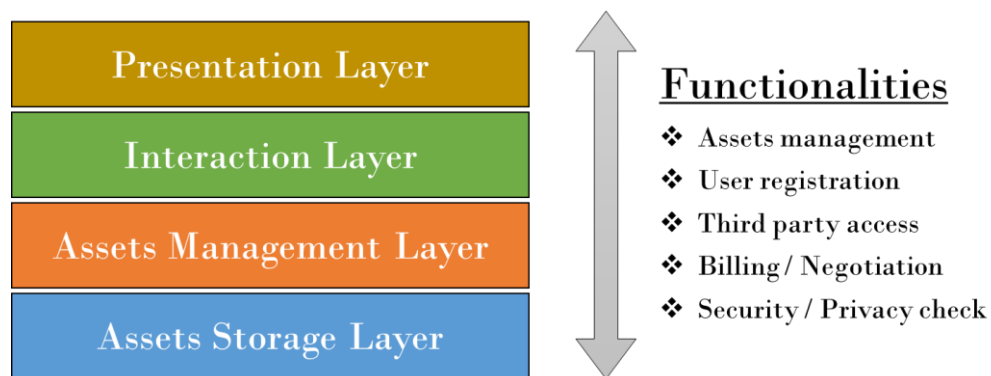

Figure 5 - Market platform's layers and main functionalities.

4.1.1. Assets Storage Layer

The Assets Storage Layer is responsible for storing the assets / objects that will be offered by the market platform (Section 2.3.1 - Offerings). An essential component of it is a database that can store files in any format as well as additional information about the files provided. For this purpose, the appropriate database type is NoSQL technology. NoSQL databases ("Not only SQL") are non-tabular and store data differently than relational tables do. NoSQL databases come in a variety of types based on their data model. The main types are document, key-value, wide-column, and graph databases. They provide flexible schemas and scale easily with large amounts of data and high user loads.

Based on the above, a NoSQL database is used for the market platform and in particular, a document-based database so that it can store files. One of the most performant and efficient document databases is MongoDB. MongoDB stores data in flexible, JSON-like documents, meaning fields can vary from document to document and data structure can be changed over time.

In addition, GridFS - a file system / specification for storing and retrieving files that exceed the maximum size limit that MongoDB sets - will be used to store files and assets divided into parts, called chunks, storing each chunk as a separate document. The documents will be stored with their binary data in two collections that contain their chunks and metadata.
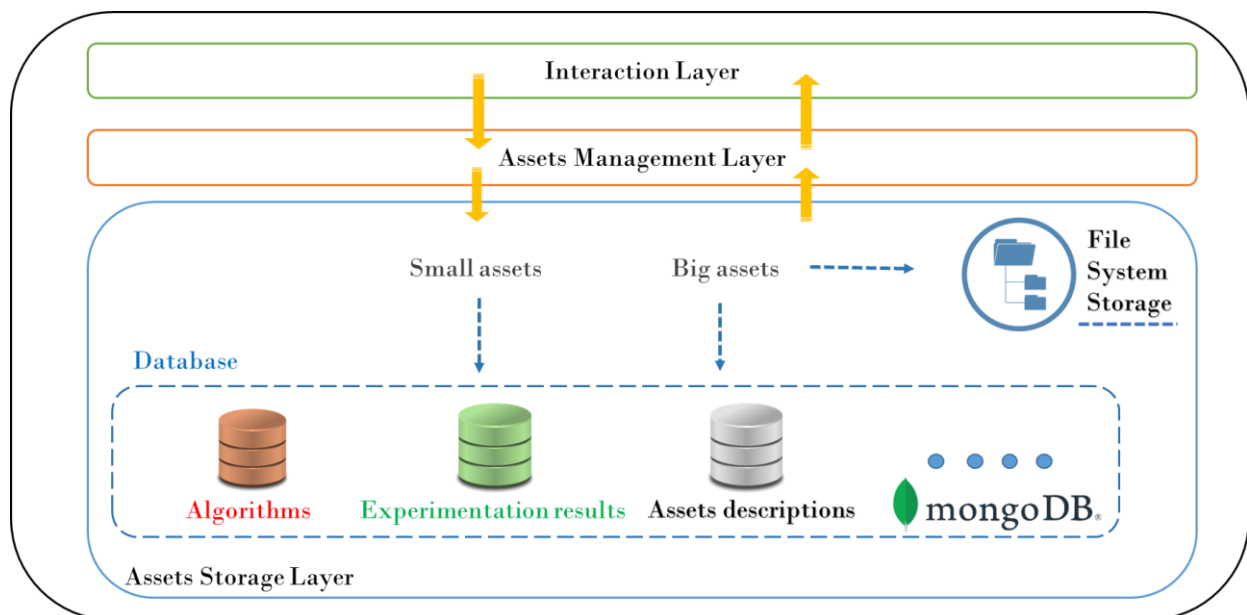

Figure 6 - The "Assets Storage Layer".

Thus, for this layer of INFINITECH's market platform, the combination of MongoDB and GridFS, is used to store small and large files. Appropriate collections have been created in the MongoDB database per destination branch / field of application of the files (i.e., a collection for VDIH services, another collection for Big Data solutions, etc.).

The information provided about the offered assets by their providers, called descriptions (described in Section 4.1.2), will also be stored in the MongoDB in JSON format, which is fully supported by MongoDB, as mentioned before. Moreover, JSON files are easy to be created, retrieved and read even by a simple user, although in the presentation layer, the containing information will be presented to the end users in an even more friendly way. JSON files also extend the interoperability of systems, especially in this case, the main core of INFINITECH's market platform and the Front-end.

4.1.2. Assets Management Layer

This layer is responsible for the entire life cycle of the data and assets within the platform and offers all the principles and techniques for their management. Specifically, the Assets Management Layer is the layer that handles the assets from the moment they are entered to the platform through the APIs and then stored in the database (in Assets Storage Layer), until their final deletion from the platform.

In general, through this layer, the market platform will support the CRUD operations. CRUD stands for "Create", "Read" ("Retrieve" in this case), "Update" and "Delete" which are the four basic functions of persistent storage. Except for these operations, the platform will also support the searching functionality, as described below:

- Create functionality, is the functionality where new assets / objects are ingested in the marketplace. This operation is triggered following the upload of a new asset by a provider and it results to the creation of appropriate entries in the database for the new document (the document in chunks, its metadata, etc.). An important information for each object refers to a unique alphanumeric that will be its identifier through which the indexing can be done as well as its subsequent retrieval.
- Search functionality, is the functionality that will enable the end-users to search for assets, based on various parameters from metadata given by the providers or possibly other information generated by the system during its creation (e.g., identifier - ID).

- ▪ <u>Retrieve functionality</u>, is the functionality that will be executed when an end user wants to download an asset from the marketplace, after first seeing / finding it through the search. The retrieval will result to obtaining the corresponding asset from the Storage Layer and delivering it to the appropriate API.
- ▪ <u>Update functionality</u> is the functionality that will be triggered when an end user replaces an asset with a newer version of it or modifies some of its metadata.
- ▪ <u>Delete functionality</u>, is the functionality that will handle the deletion of the assets from the platform when the owners, or other users with the appropriate rights, decide to delete them.
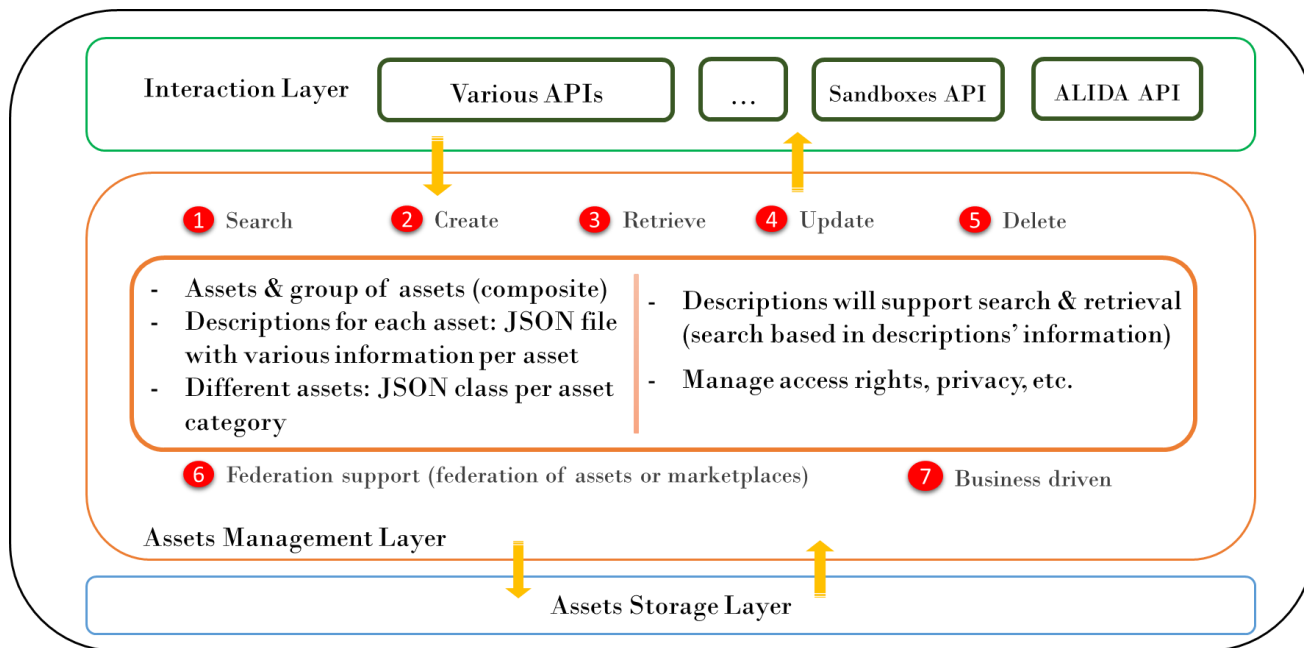
Figure 7 - The "Assets Management Layer".

The main part of the marketplace, i.e., all layers except the Presentation Layer (because of its separation from the backend), is managed by RESTful APIs. A RESTful API is an application program interface (API) that uses HTTP requests to GET, PUT, POST and DELETE data / assets. These words in capital letters, are some HTTP methods which are related to the previous Search and CRUD operations / functionalities. The following table matches these HTTP methods with the above supported functionalities and the corresponding statements in MongoDB (the database that will be used):

Table 4 - Matching (S)CRUD operations with HTTP methods and MongoDB statements.

Operation	HTTP method	MongoDB
Create	POST / PUT	Insert
Retrieve	GET	Find
Update	PUT / POST	Update
Delete	DELETE	Remove
Search	GET	Find

An important feature of the marketplace is the search functionality. When a provider uploads a new asset to the marketplace, he / she will be prompted to provide additional information about the asset (i.e., metadata). This information could be related to the usability of the file, the type of the file / asset, its input parameters as well as the exported results if it is an algorithm, who has the right to access it, other useful comments, etc.

This set of metadata related to an asset, will be the content of a JSON file called "**description**" file of the asset. Each asset will have a description file so that it can be searched and retrieved by end users, who will

be able to search for files according to this information, i.e., metadata. Therefore, the submission of these description files per asset is necessary because otherwise assets' existence on the platform will not make sense and they will probably not be able to be retrieved.

It will also be possible to obtain the assets' information and automatically convert them to JSON files through the front-end, but in cases where the end users interact directly with the APIs of the Interaction Layer (e.g., the INFINITECH sandboxes), they will need to create on their own, appropriate description file. Based on the above, as the offerings vary and since different information may be needed for each type of offered asset, templates with minimum required fields per asset type / class are created (the final description templates / models will be delivered in the next WP8 deliverables). A representative example of a description file for K-means algorithm in JSON format is cited below:

```
{
    "id": {"id":"fjsi48hfsf057-43fjd", description": "unique identifier"},
    "owner": "sample user",
    "properties":{
        "type": "ML Algorithm",
        "name": "Kmeans",
        "task": "Clustering",
        "description": "Entity's basic properties"
    },
    "executable_details":{
        "filename": "kmeans.py",
        "runtime_environment": "Python",
        "libraries_required": { ScikitLearn": 0.22 },
        "description": "environment and libraries required for execution"
    },
    "input_parameters": {…},
    "output": […],
    "comments": […],
    "other_details":{
        "author":"sample user",
        "created_at":"06.05.2020",
        "last_updated_by":"sample user",
        "updated_at":"10.06.2020"
    }
}
```

4.1.3. Interaction Layer

This section describes the different aspects and interfaces of the RESTful APIs that will handle the market platform operations. More specifically, it describes the envisioned interfaces that will connect the outside world to the platform. The marketplace will have different APIs that will allow all users of the system, providers and end users, to interact with the platform.

As for end-users, APIs will be developed and will receive HTTP requests for retrieving the offered assets but also information that will enhance the search functionality (Section 4.1.2). For example, these APIs will enable users to search the description files and / or metadata stored in the database to find information that fits the needs of end users. Of course, since end-users interact with the market platform through the front-end, an API is required to enable the front-end and the back-end to exchange HTTP requests and / or information in JSON format.

On the other hand, other systems and components, such as the INFINITECH sandboxes will have at least one (depending on the needs) different API that will offer them more direct access. For example, the overall architecture of ALIDA platform (https://alida.alidalab.it), a micro-service oriented platform for the composition, deployment and execution of Big Data Analytics (BDA) applications, contains a "Service Catalogue" component, with which the INFINITECH marketplace could interact (via an API) in order to provide to the ALIDA platform descriptions and executables from the marketplace (or vice versa).

The same applies to the providers as users of the market platform. Most of them will interact with the platform through the intermediate component, for example the front-end, which will send HTTP POST

requests (using the same API in which it sends the rest HTTP requests) in order to store new assets in the marketplace.

As described in the previous section (Section 4.1.2), there will be different APIs per asset type, in the same way with the collections of the database. In this case, the platform's users (systems and humans) will share the same APIs in order to send their HTTP requests and consequently there will not be APIs for each type of user (providers and end users).
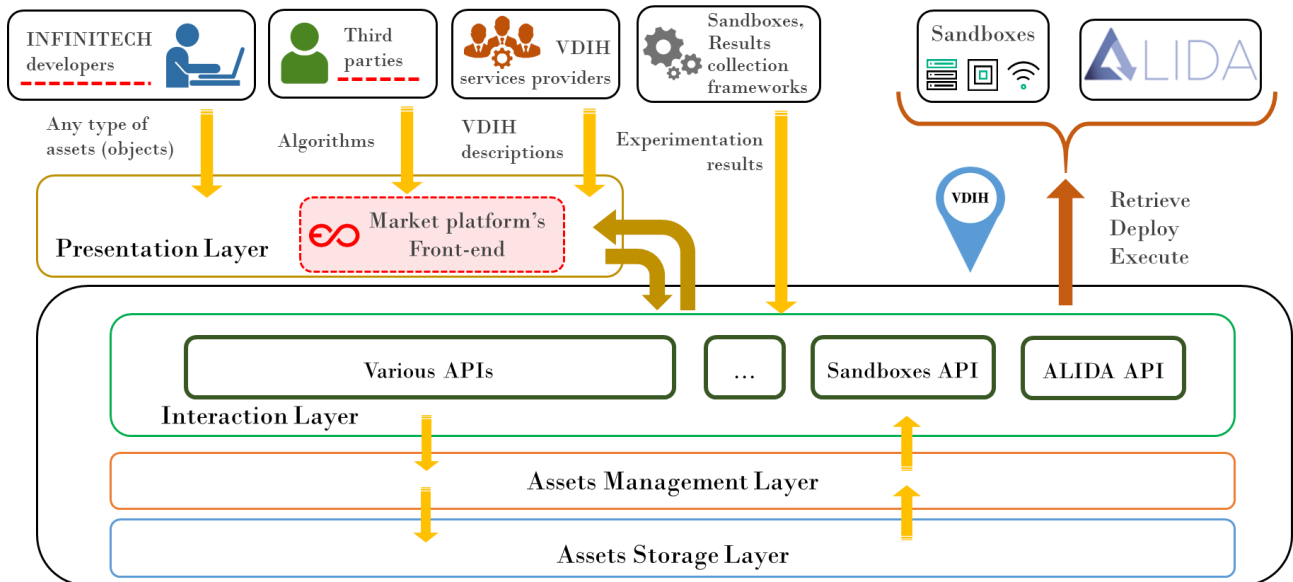

Figure 8 - The "Interaction Layer".

In addition to the ALIDA platform which has been mentioned earlier, the market platform could also be linked to other important platforms and repositories for data scientists, with emphasis on resources for digital finance, and to libraries and tools that are used by INFINITECH. Here are some examples:

- OpenML (https://www.openml.org), an open science platform with datasets, algorithms and tasks used for machine learning (ML). It provides important benefits for the science community and beyond.
- MLBox (https://mlbox.readthedocs.io), a powerful automated machine learning python library that provides state-of-the-art solutions for machine learning.
- Kaggle (https://www.kaggle.com), a platform intended for the data science community, offering tools and resources (e.g., datasets). INFINITECH's marketplace could retrieve open financial datasets, such as credit card datasets, fraud detection datasets, alternative lending etc.
- Data.gov (https://www.data.gov) – A Catalog of over 200,000 open-sourced government datasets in topics such as agriculture, climate, consumer, ecosystems, education, energy, finance, health, local government, manufacturing, maritime, ocean, public safety, and science.
- Google AI Platform (https://cloud.google.com/ai-platform), a managed service that enables the easy building of machine learning models, that work on any type of data, of any size. Periodically, it releases various data of interest, including financial data.
- KNIME Hub (https://hub.knime.com), also described in Section 1.3.2, is an open tool for data-driven innovation, designed for discovering the potential hidden in data, mining for fresh insights, or predicting new futures.
- Registry of Open Data on AWS (RODA - https://registry.opendata.aws), which is a registry containing collections of datasets that are publicly available through Amazon Web Services (AWS). These datasets are not provided or maintained by AWS, but by third parties, so licenses will be verified in order to retrieve / use them.
- UC Irvine Machine Learning Repository (https://archive.ics.uci.edu), which is also a machine learning repository, containing a diverse collection of over 400 datasets including some finance related.

In the cases where it is possible to use an API for more than one system, then this "shared use" may be preferred over the creation of new same APIs / portals for the respective systems.

Finally, and based on the proposed specifications for the back-end and the front-end, the respective domains have been realized offering their APIs for them to be accessed, which are the following:

Back-end: https://db.infinitech-h2020.eu

Front-end: https://marketplace.infinitech-h2020.eu

4.1.4. Presentation Layer

The following figure (Figure 9) depicts the interactions between the front-end / back-end and the additional components of the overall INFINITECH market platform.
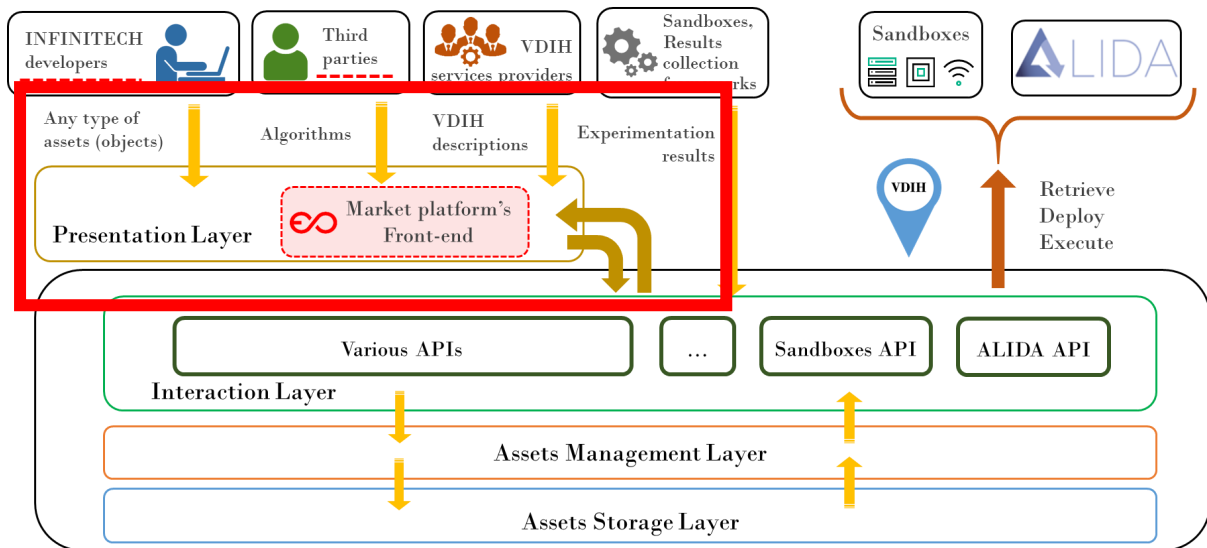
Figure 9 - The "Presentation Layer" and the front-end (within the red box).

The Presentation layer was developed (initial version) taking into consideration the kind of information to be presented.

A structure has been defined for the marketplace, present in the following image:
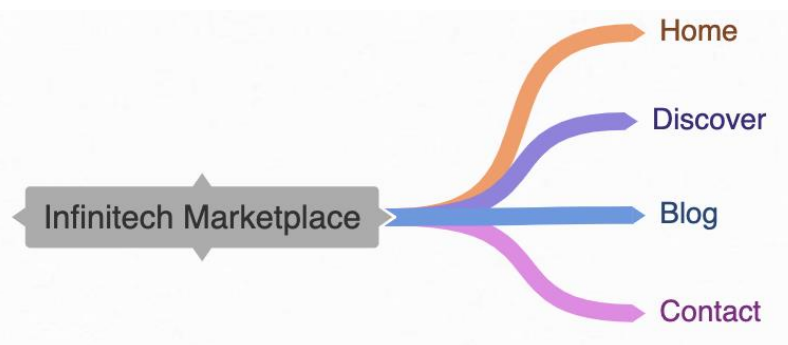
Figure 10 - INFINITECH Marketplace mindmap

Home page

The pages are immediately available through the navigation bar, included with every page, it also has a search functionality. The home page includes a big button which redirects to the Discover page, and three stats that reflect the total amount of INFINITECH Resources, the total amount of assets and the total amount of support tools. We then get a list with four examples of *assets or support tools*, and a list of the latest training activities. The footer is also included in every page, providing some information about the project and navigation.

Figure 11 - Home page

Discover Page

At the Discover page we are presented with a solution for navigating through all the *assets/support tools* provided by INFINITECH's consortium along with pagination and buttons for filtering results through categories. Each results card contains a title, a description and links to the tags and categories pages. A side menu is also present which contains a search widget and a list of the latest assets (and support tools).
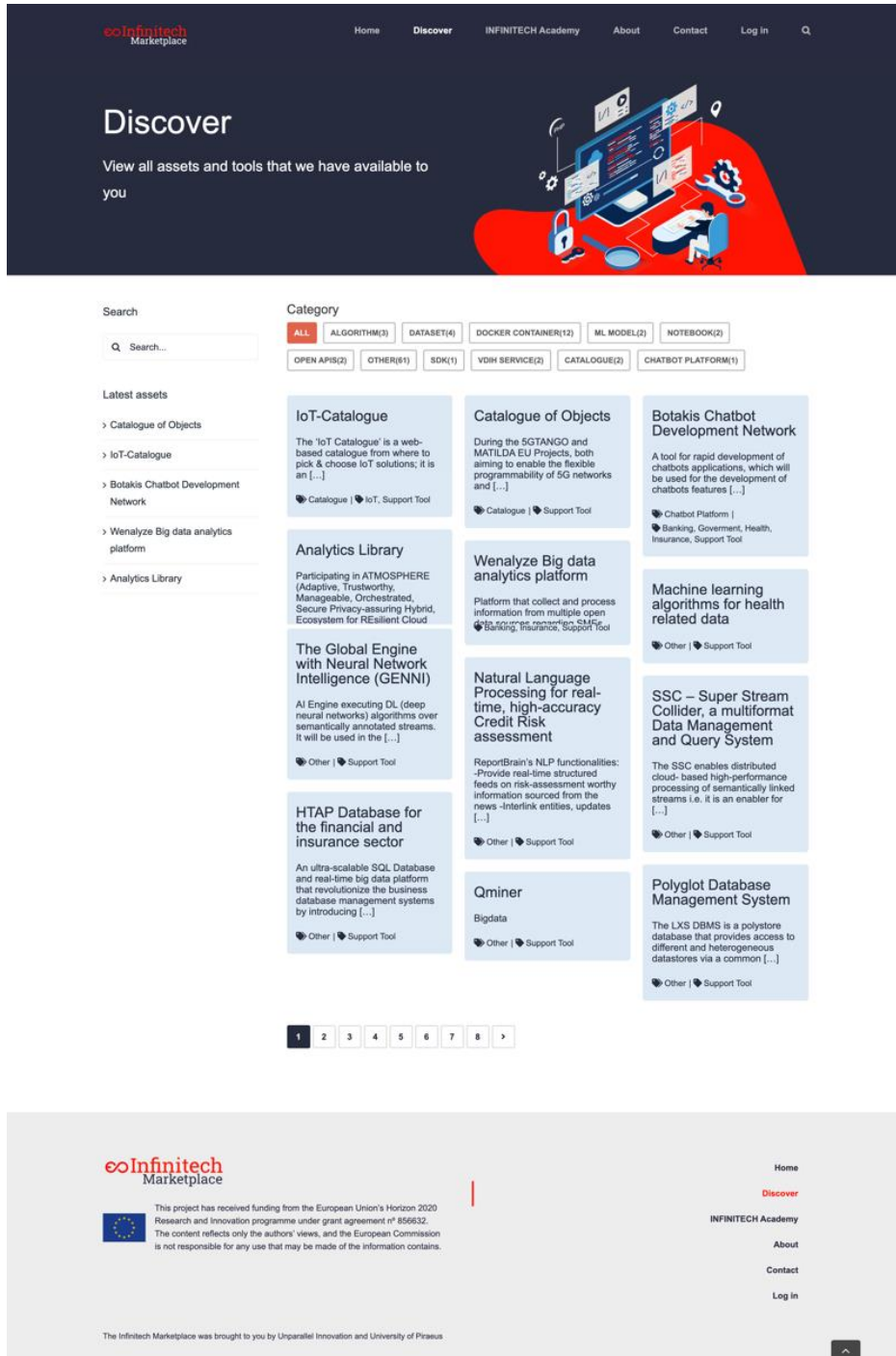
Figure 12 - Discover Page

Asset or Support Tool Page

Through clicking in one of the titles in the cards we were previously presented with, we arrive at a page which displays information about the Asset or Support Tool. It shows the following information, in addition to the asset title and a description:

- Owner – the owner of the asset or support tool
- Estimated Release Date – estimated release date
- Asset Composition – either "Simple" if it is a single thing or "Composite" if it is a combination of various tools
- Link – displays link if available
- Category – category of the asset or support tool
- Type – either "Asset" or "Support Tool"
- Field of use – field where the technology is used

Every Asset or Support Tool has a page like the following.
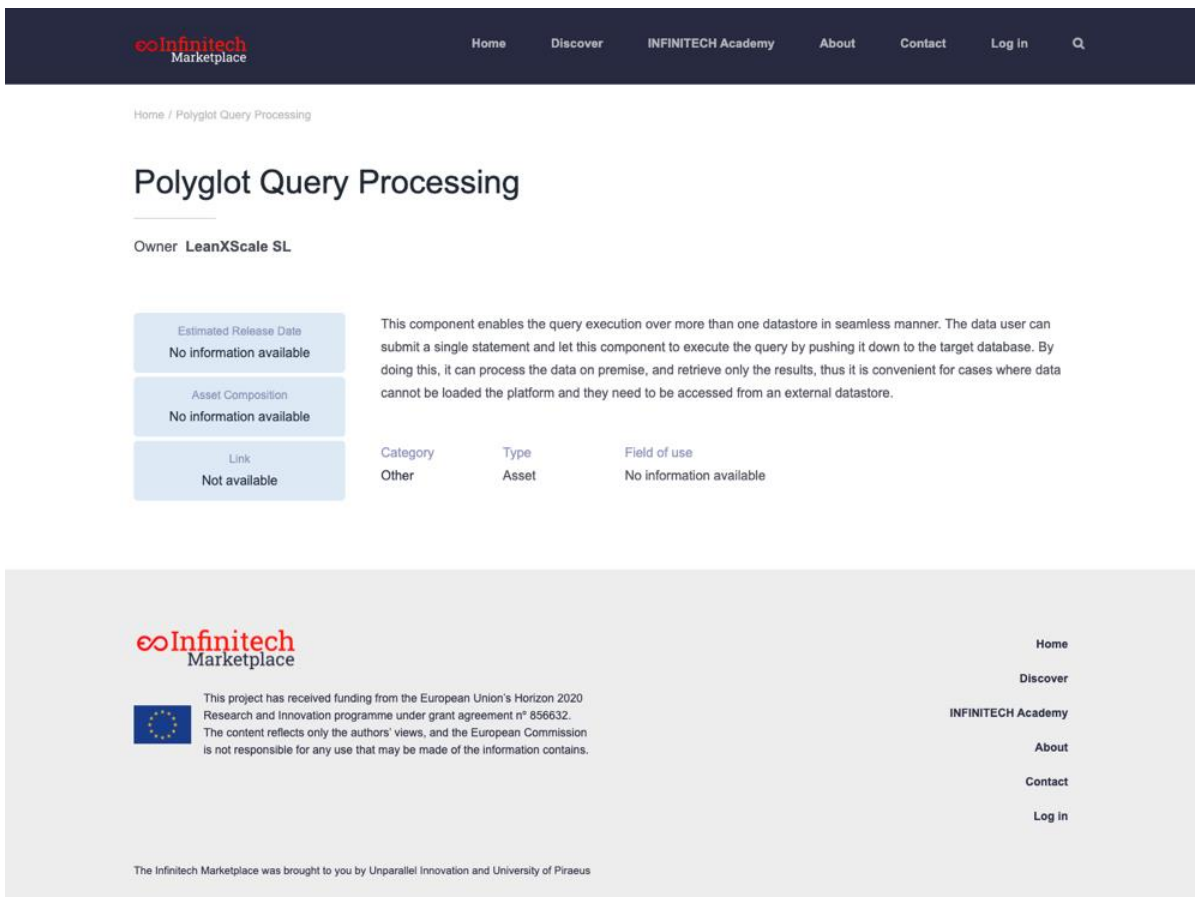
Figure 13 - Asset or Support Tool Page

INFINITECH Academy page

Gives us access to various workshops in video format. Once you click one of the items you get access to a video and more information about the workshop.
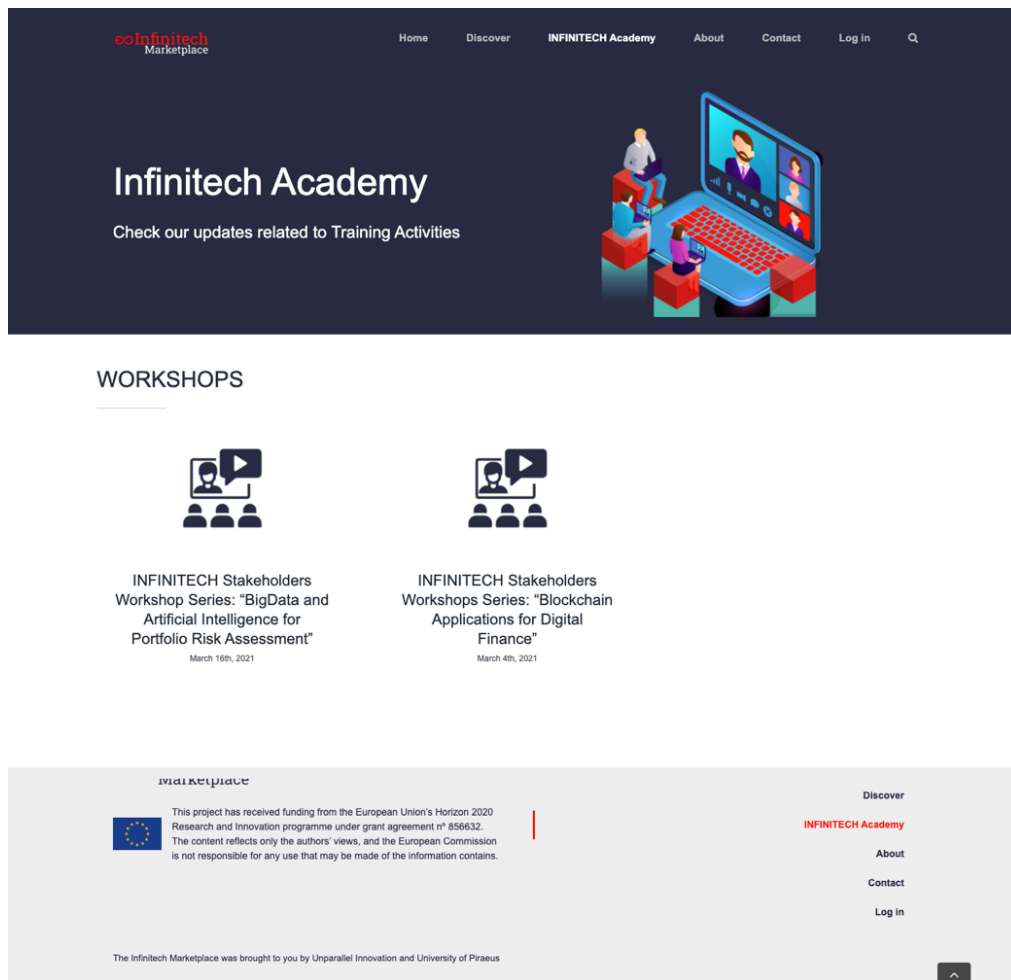
Figure 14 - INFINITECH Academy Page

Figure 14-b - INFINITECH Academy Page 2

About Page

In the About page we highlight the connection of the marketplace with the INFINITECH Project. By clicking in the "Visit Here" button, the user is redirected to the main Project website.
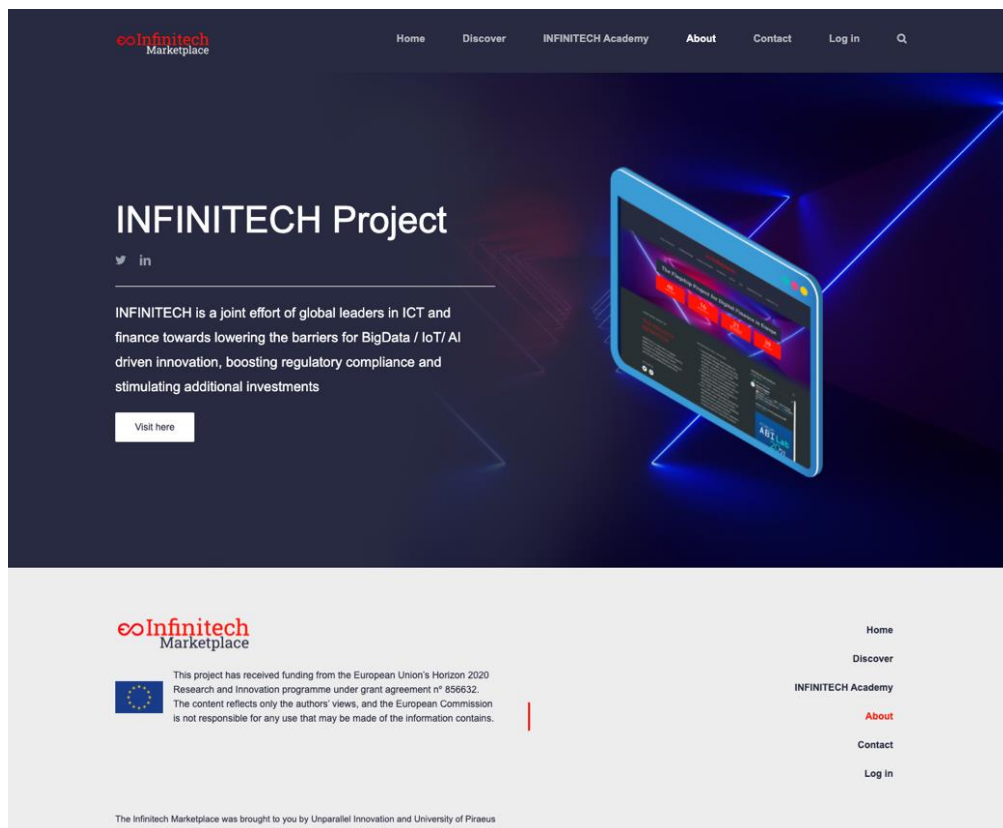

Figure 15 - About Page

Contact Page

The Contact page presents us with a form as a way to send a message to the INFINITECH consortium, for any questions stakeholders might have due to their interest in the marketplace or in any specific resource available in the marketplace.
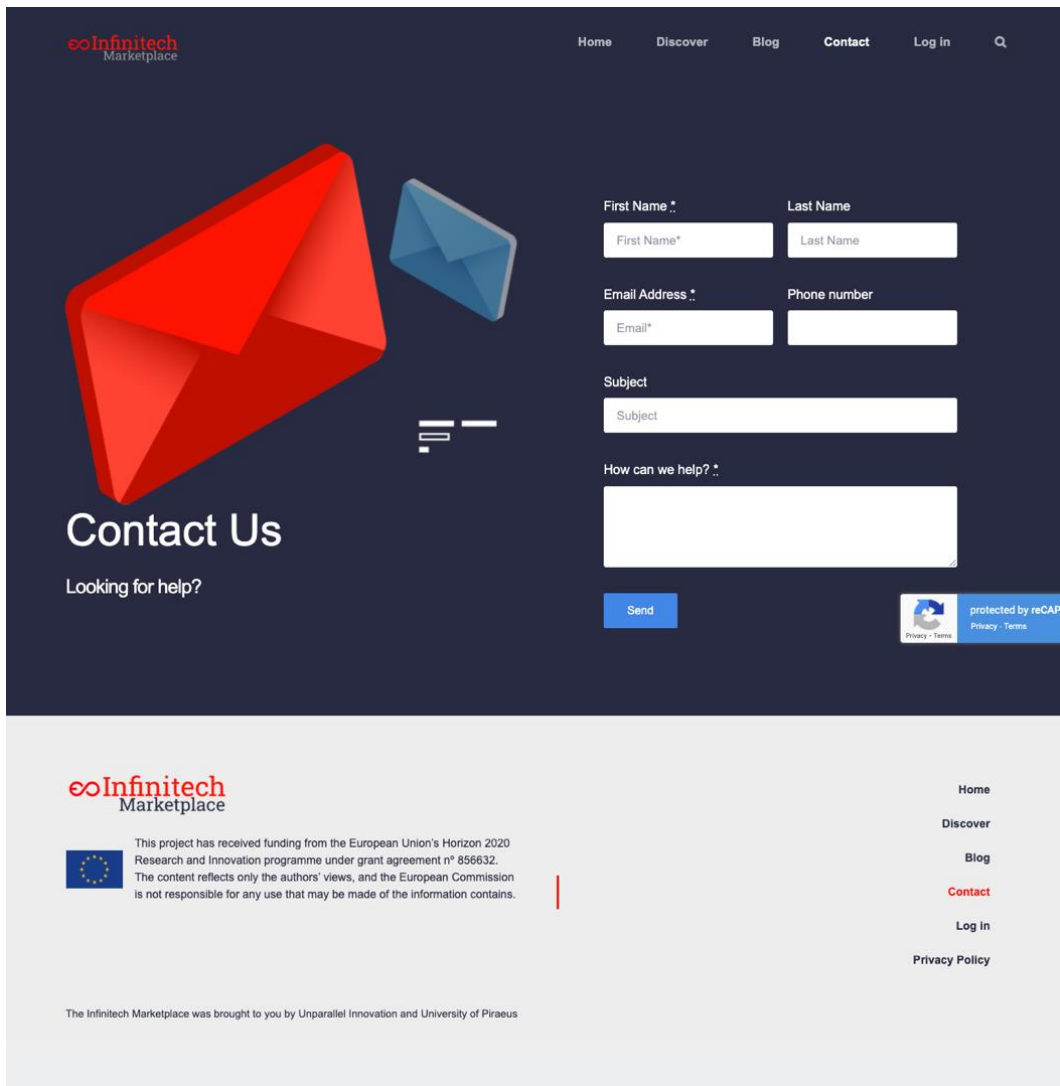

Figure 16 - Contact Page

4.2. VDIH Specifications

4.2.1 Services Overview

The Virtualized Digital Innovation Hub (VDIH) is a component of the INFINITECH marketplace and will be integrated in it to provide innovative services ensuring a complete set of offers that can be presented to customers and businesses. The VDIH's services will leverage on the expertise of the consortium and will be differentiated in categories, providing filters for agile research and identification. The VDIH will be based on the INFINITECH marketplace infrastructure, and its dashboard will be a specific layer inside the Presentation Layer.

The services provided by the VDIH can be categorized in two main groups:

- **Digital Innovation Services (DIS)**. These include the usual services provided by the Digital Innovation Hub but in a totally Virtualized manner. The DIS are a very relevant layer of the VDIH and they are presented through an easy and user-friendly dashboard that uses filters and maps to give customers an effective path for finding the right service. The partners of the INFINITECH consortium will provide joint expertise and knowledge in different fields, offering support to companies to become more competitive with regards to their business processes, products or services using digital technologies. Some already identified services are:
  - **Technological support services.** Such services leverage on the technological expertise of the consortium for providing a vast range of services such as engineering, design, specification and implementation. Third-parties outside INFINITECH will be offered with the option of providing requests for technological support (e.g., "Help me with my FinTech solution or MVP) through the INFINITECH Marketplace.
  - **Training services**. Several training services would be provided to enhance the businesses' employee's level. The training may vary from technical to quality or skill training. In this direction, the project will provide a range of training materials and courses through the marketplace, covering topics related AI, Blockchain and BigData in digital finance.
  - **Business planning services**. These services would aim to implement strategic approaches to merge innovation and business exploiting cutting-edge technologies and skilful collaborators. As per the context of INFINITECH, such planning would be much targeted to the financial environment. Third-parties outside INFINITECH will be offered with the option of providing requests for business planning services (e.g., "Create a business plan for my Business Case or FinTech company).
  - **Business modelling services.** Innovative business model services would be provided to translate complex processes into financial insights and decision-making approaches. The existing framework, such as Business Model Canvas or Odyssey 3.14, would be utilized as a means to achieve strategic models and approaches. Third-parties outside INFINITECH will be offered with the option of providing requests for business modelling services (e.g., "Propose and help me validate a business model for my Business Case or FinTech company).
  - **Go-to-market services.** These would take into account the essential factors to develop effective strategies, aimed at increasing the technology readiness level (TRL) of a specific product, providing the most fruitful market segmentation for a business, or defining the target market for a service.
- **Sandboxes Experimenting Services (SES).** This second category of services is mainly devoted for providing design services, safe testing environments and innovative joint labs models. The infrastructure developed in INFINITECH would allow testing and deployment of solutions in safe environments - sandboxes.
  - **Co-design services.** Companies and businesses would be guided in the overall design of a specific product or solution in gathering requirements, defining specifications, developing the architecture, selecting technical services and then being supported for the integration.
  - **Testing services.** The consortium would offer safe environments to allow stakeholders testing their solutions in an easy and friendly way, leveraging on the flexibility of the INFINITECH Reference Architecture and its microservices approach. Such environments may be both for early-stage testing and proof of concepts, and for real testing environments and prototyping.
  - **Co-innovation services.** These services would leverage on the available innovation labs provided by INFINITECH consortium, supporting the businesses in creating new innovative models, enabling them to access the market through joint collaborations.

The VDIH would be the catalogue of the services stated above and more. These will be appropriately structured and their delivery will be defined based on the federation and integration of resources from existing innovation management infrastructures of the partners. High importance would be given to access and governance rules for ensuring easy-to-use and fully compliant services.

4.2.2   Related Assets

The services of the VDIH will be supported by assets of the INFINITECH market platform, including the technological assets of the platform and complementary assets like training. Specifically, the following enhancements to the already described assets will be used to support the VDIH and the delivery of its services:

- **Packaging of Assets**: To boost the technical support, integration and sandboxing services, some of the INFINITECH assets will be packaged and made available as Docker containers. This will facilitate their distribution, deployment and use in different environments.
- **Technical Documentation of Technological Assets**: Some of the technological assets of the marketplace will be appropriately documented to facilitate their use in technological support and sandboxing services. This documentation shall include:
  - **Developers' manual,** including documentation of APIs, sample code and how-to guides.
  - **Users' manual,** with information about potential use of the asset in digital finance scenarios.
  - **Demonstration of the asset,** through screenshots, videos and simulation.
- **Training and Education Assets**: To support the training services of the VDIH the assets of the marketplace will include:
  - **Presentations and tutorials** on Digital Finance topics.
  - **Structured Courses** in AI and Blockchain in Digital Finance**.**
  - **Webinars,** including recorded workshops of the INFINITECH projects**.**
  - **White Papers** and Publications of the project**.**
  - **Links to external resources,** including resources residing within affiliated platforms and marketplaces (e.g., the IoT Catalogue and the Finsecurity.eu platform of the FINSEC project).

4.3.   Overall Conceptual View

Following the description of the individual elements, the overall conceptual view of the INFINITECH market platform is presented in the following figure (Figure 17), which depicts all platform layers along with their key offered functionalities, the providers and end users.
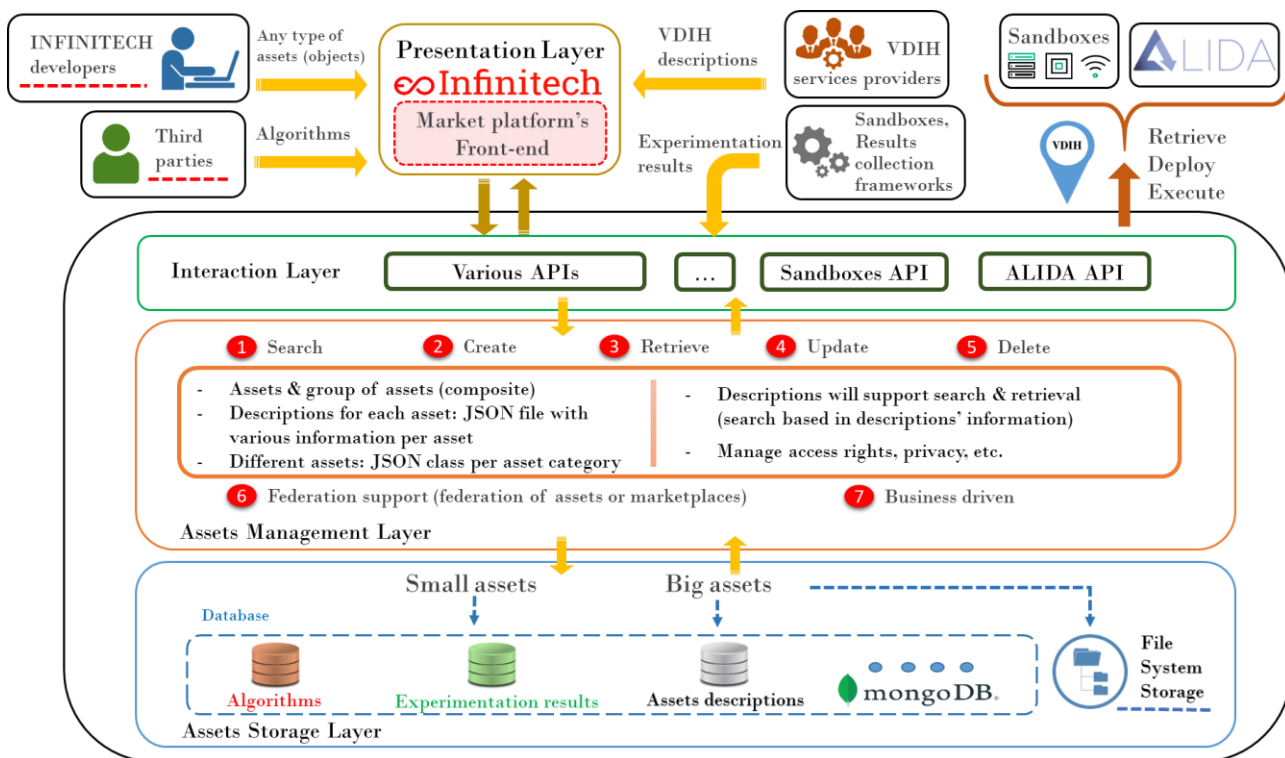

Figure 17 - Overall architecture.

The overall information flows are depicted in the figure through the respective arrows that represent the main interactions:

- The four-perspective hybrid market platform will interact with its users through the front-end (Presentation Layer) from which HTTP requests will be sent to the platform (requests depending on the case: search, create / store new assets, retrieve an asset, update or delete an existing asset).
- These HTTP requests will be received by the corresponding APIs of the Interaction Layer. Users that can also be systems will be able to interact directly with the APIs.
- After receiving the requests, the Assets Management Layer will undertake the processing of the requests, using the developed functionalities. Specifically, it will interact with Assets Storage Layer to retrieve useful information that, after processing, will be sent to the users via APIs, in response to their HTTP requests.
- When a provider intends to upload a new asset, the provider will also submit the asset description file (via front-end it is generated automatically by filling in the appropriate fields), which will contain metadata for the asset. Both the asset and the description will be stored in Assets Storage Layer, in the database (MongoDB) if they are small files, otherwise in the file system (GridFS).
- In order to retrieve the assets, the Assets Management Layer, through the retrieve functionality, will find the requested asset from the database and will deliver it to the end user via the platform's APIs. The retrieval of the assets will be also done in a similar way by the sandboxes and the linked platforms, repositories, libraries, etc.

In the following subsections, additional core use cases are introduced to highlight the interaction of the end-users with the market platform.

4.3.1. Use case: Upload assets

The following figure presents the use case of uploading a new asset in the marketplace. Note that in the figure, the area with the yellow background cites the case where users upload the assets (and their descriptions) through the front-end, while the area with blue-gray background reflects the information with respect to the back-end interactions. The IDs (either for assets or descriptions) refer to unique IDs that are associated to each asset / description. As described in section 4.1.2, the models for the description files that are mentioned also below, will be delivered in the next WP8 deliverables.
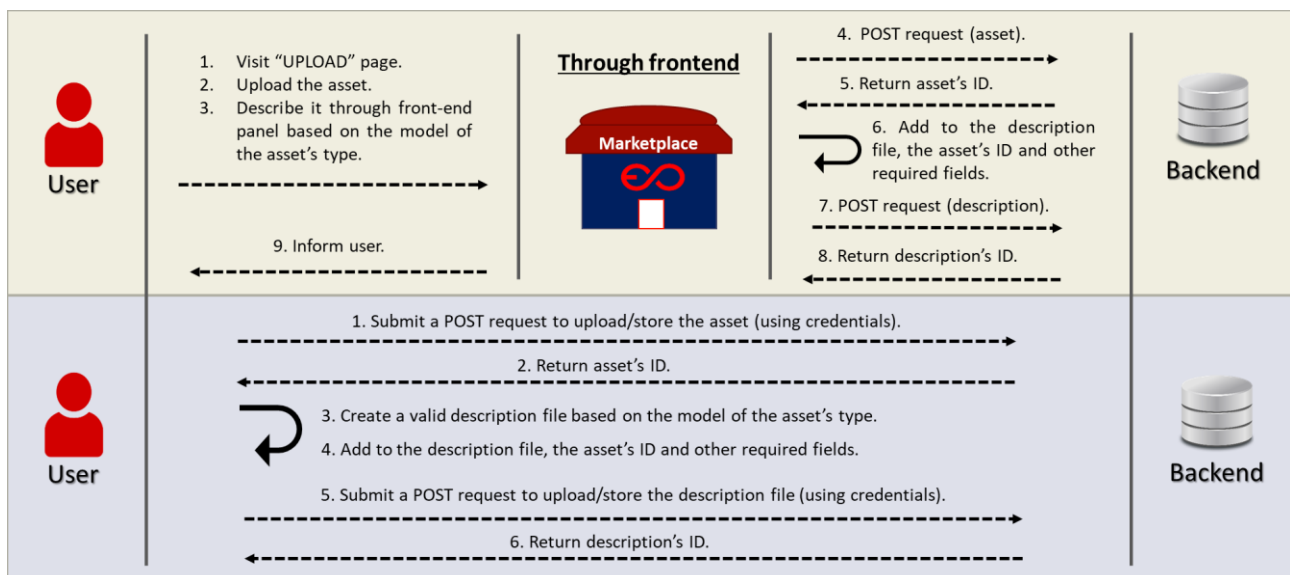

Figure 18 - Upload use case

4.3.2. Use case: Update assets

The update use case has a similar process with the upload use case (section 4.3.1), with the difference that in this case, the assets should be already stored in the market platform. Corresponding to the previous use case, Figure 19 depicts the interactions in terms of front-end and of the back-end of the market platform. Status messages, a point to be elaborated regarding the information on the figure, refers to HTTP status messages that back-end server responds to clients' requests, to inform them regarding the requests' process.
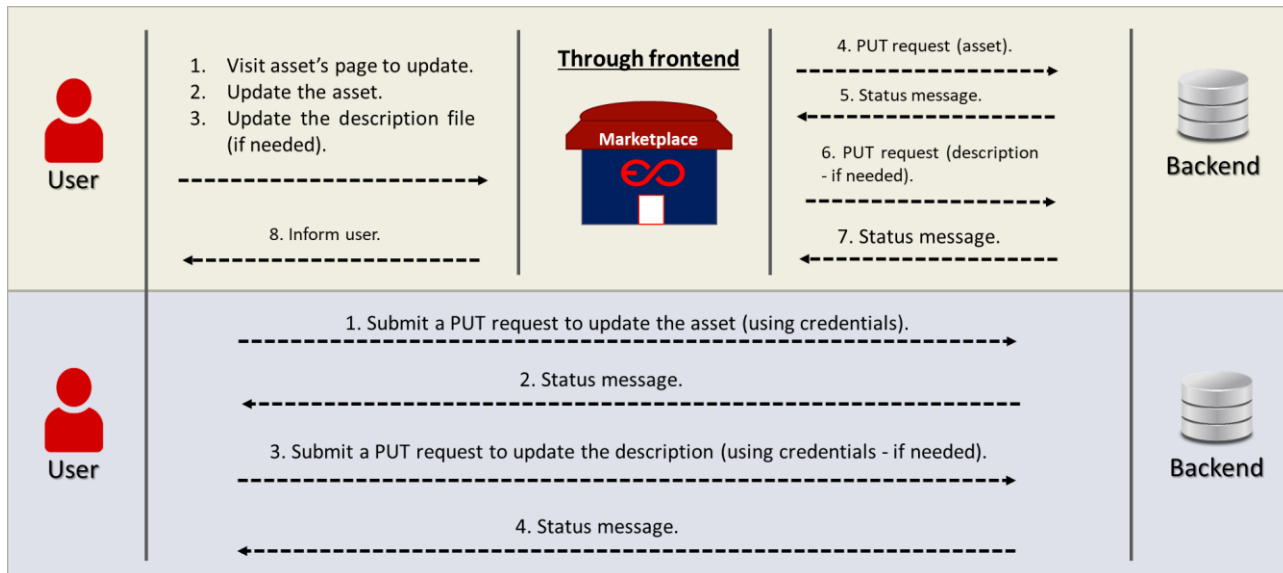

Figure 19 - Update use case (through Front-end).

4.3.3. Use case: Retrieve assets

The following figure (Figure 20) depicts the process of retrieving an asset from the marketplace, either through the front-end or directly through the back-end. As depicted below, the descriptions that are downloaded refer to the descriptions as the one presented in section 4.1.2. Another point that needs to be further extended for the following figure (but also for the previous figures) are the credentials. As described in section 4.1, the platform will take into account issues related to security of the assets and thus the access of the assets will be done through credentials, after user registration and / or specific rules that may had been provided by the owner / provider of the asset.
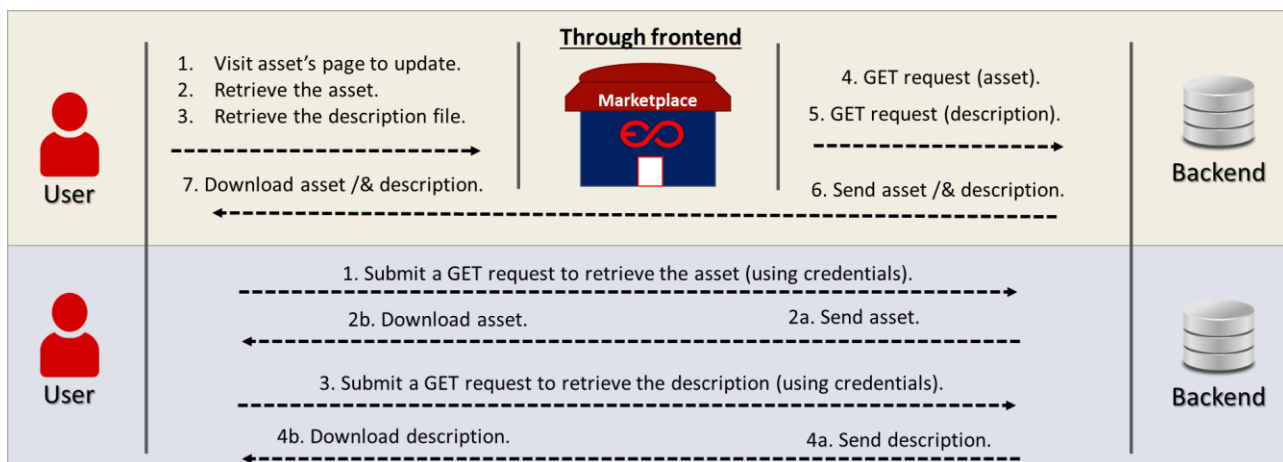

Figure 20 - Retrieve use case.

4.3.4. Use case: Delete assets

The process for deleting assets is presented in Figure 21 for the front-end and back-end services of the market platform. This process will be available only for the owners of the assets and for the moderators of the market platform. As in the update use case, the status messages that are reflected in the below figure, are HTTP messages that will inform the users whether their assets have been successfully deleted or not.
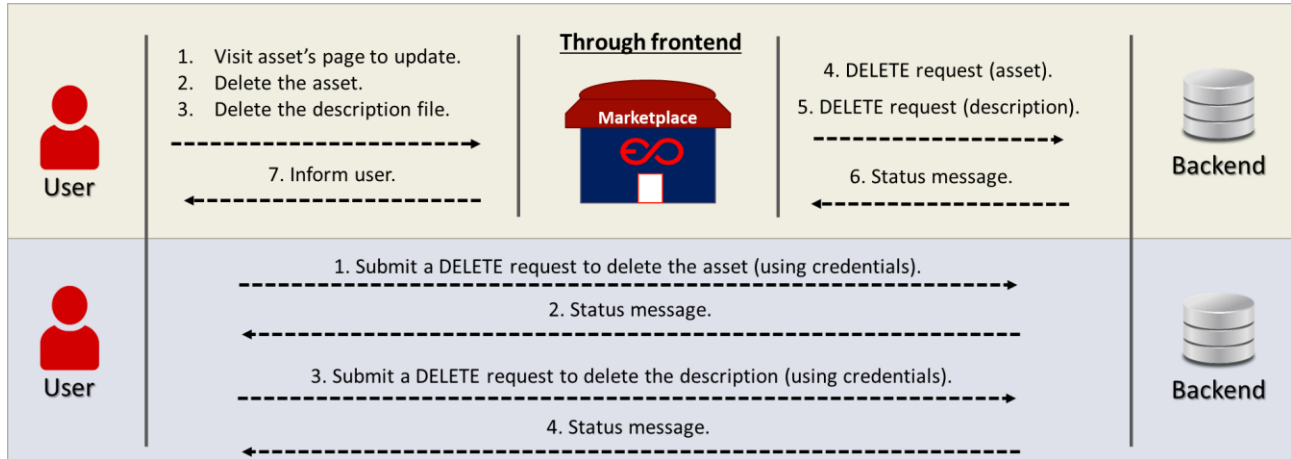

Figure 21 - Delete use case.

4.3.5. Use case: Search assets

One of the main capabilities of the market platform refers to the search functionality. The figure below cites the corresponding information flows for the back-end and the front-end. The search functionality is applied to various metadata that are the content of the description files (as described in section 4.1.2). Thus, the search functionality is applied directly to the descriptions and this is the reason that the following figure presents the descriptions in parentheses. As depicted below, search parameters / filters are set in the URL as query strings (GET request).
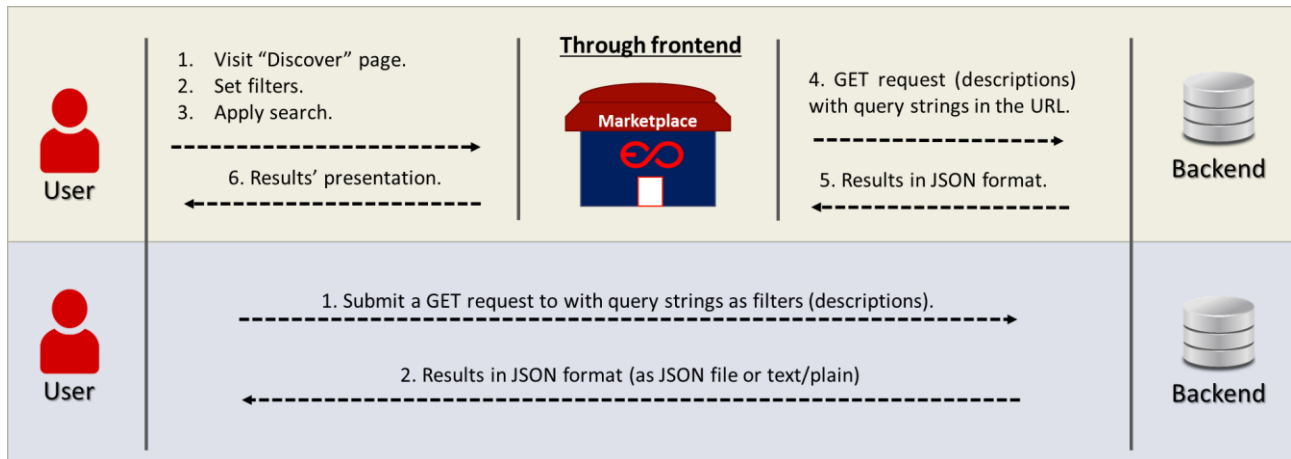

Figure 22 - Search use case.

5. Conclusions

This report provides the updated specifications of the INFINITECH market platform. The specifications are utilized for the realization and implementation of the market platform, encompassing the main functionalities (reflected in the respective main layers of the market platform that have been introduced in this deliverable) regarding the storage of assets, the advanced search and retrieval, and their combination into unique turn-key offerings.

Overall, the INFINITECH market platform will integrate ready-to-use solutions and assets of the project, including analytics algorithms and analytics pipelines (i.e., workflows of algorithms), datasets / data assets, experimentation results of the aforementioned algorithms as well as validated turnkey solutions for finance and insurance. The market platform will also act as a digital innovation hub by hosting innovation management services, the so-called Virtualized Digital Innovation Hub - VDIH services that can be exploited by FinTech and InsuranceTech practitioners.

Based on these specifications, the initial version of the market platform provides a single-entry point for accessing resources for big data, IoT and AI innovations in the finance / insurance sectors, being also the main enabler for the INFINITECH exploitation strategy. To this end, the market platform is implemented and populated with assets obtained both from 3rd parties and from INFINITECH researchers (e.g., algorithms and datasets). As an initial version has already been made available (https://marketplace.infinitech-h2020.eu), during the next phase of the project the main focus will be on extending the current version with additional functionalities (e.g. in terms of correlated assets) as well as populating it with information that emerges from the project and from 3rd party contributors. The planned WP8 deliverables (D8.3 - BigData and AI Solutions, D8.5 - IoT and Blockchain Solutions, D8.7 - Virtualized Digital Innovation Hub, D8.9 - Third-Party FinTech and InsuranceTech Solutions) will detail the respective outcomes of the performed activities.