



Specifications
for the ARIANE project

Summary

I - Contexte	2
II - Needs, constraints and expectations of Prometheus	3
III - Software specifications	4
3.1 Overall software design	4
3.1.1 The 2 purposes of the software	4
3.1.2 Software modules	5
3.1.2 The Pivot Ontology and authority frameworks	6
3.1. Functional specifications of the modules to be produced	7
3.1.0 "Pivot Ontology Management" module	7
3.1.1 "Edtech Management" module	7
3.1.2 "Management of Training Organizations" module	10
3.1.3 "API Ontology" module	10
3.1.4 "API Structures" module	10
3.1.5 "Data API" module	10
3.2. Ergonomic specifications	12
3.3. Technical specifications	13
3.3.1 Main technical components	13
3.3.2 Production chain	14
IV - Delivery times	16
V - Contributions in source code	17

I - Contexte

The Association **Prometheus** aims to develop and organize infrastructure services for the circulation of education and skills data in order to enable players in the digital educational, training and employment sector to pool and share data with a view to improve their services to personalize learning, training, guidance and professional integration, as well as to use data for research purposes.

The project **Prometheus-X** named **ARIANE** is an initiative of the working group “Data Space Education & Skills (DASES)” from the French Hub of GAIA X, which aims to enable the circulation of data in an ethical, sovereign and person-centered environment. This working group decided to support the launch of the Association in order to provide a common response to the ATF call for projects from the Public Investment Bank (BPI).

The choice of the associative form and broader governance allows the Prometheus-X project to preserve its independence from financial interests and to work for the establishment of education and training data circulation services. skills in a democratic and open manner. The services developed therefore aim to serve the common interest of the sector and to guarantee compliance with the criteria of sovereignty, openness and interoperability through a common good on which all education stakeholders, guidance, integration, training and employment will be able to build the education of today and tomorrow, throughout life, centered on the person and respectful of his rights.

On September 22, 2021, the Association responded to the BPI's "ATF" call for projects in order to finance the first data circulation services and the first use cases to launch a national and then European data and communication ecosystem. education and skills services: Data Space Education & Skills (DASES).

On December 15, 2022, the BPI selected the Prométhéus-X project and granted it funding to enable it to enter into contracts with companies with a view to developing the first services and use cases.

In this context, **Mindmatcher** has the mission of developing software “***Interoperability of skills frameworks***” which contributes to the achievement of the social purpose of the Prometheus Association while respecting the objectives of common interest of the educational digital sector and the criteria of sovereignty, openness and interoperability.

The purpose of this document, Specifications, is to describe precisely and exhaustively the scope of the software to be developed within the framework of ARIANE, including in particular:

- the needs, constraints and expectations of the Prometheus Association;
- the ergonomic, functional and technical specifications of the software, particularly with regard to the objectives of openness and interoperability;
- the delivery time.

II - Needs, constraints and expectations of Prometheus

Data on a person's skills is today fragmented between different actors (schools, public employment services, universities, training organizations, employers, edtechs, platforms, etc.). This fragmentation allows neither the person to have a full awareness of their skills nor the actors to offer them the best opportunities.

The creation of the software “**Interoperability of skills frameworks**” must allow these actors to pool data to offer better services to people (employment, training, education, etc.) thanks to an ethical and person-centered network of personal data and thus making it possible to provide training to the person and lifelong integration personalized and adapted to their needs. Each person will be able to share their data, with their consent, between the different actors to have access to these best services. Each actor will be able to access the data to improve their services.

Skills data is described in numerous frameworks which differ depending on the organization. The purpose of this module is to offer interoperability services to users/applications of the “Education & Skills Dataspace”:

- Creation of a Gaia X pivot ontology for a common model,
- Translators data from partners thanks to the pivot model,
- Management of professional skills frameworks,
- Alignment of skills frameworks with authority frameworks (ROME, ESCO, ...),
- Management of training offers by skill and skill blocks.

The expected results of this task are as follows:

- Translation module with regard to the pivot ontology,
- Module for entering a skills framework,
- AI modules for aligning a skills framework with ROME and ESCO,
- Module for entering a training offer,
- AI skill suggestion modules and skill blocks,
- API for exchanges with these modules.

III - Software specifications

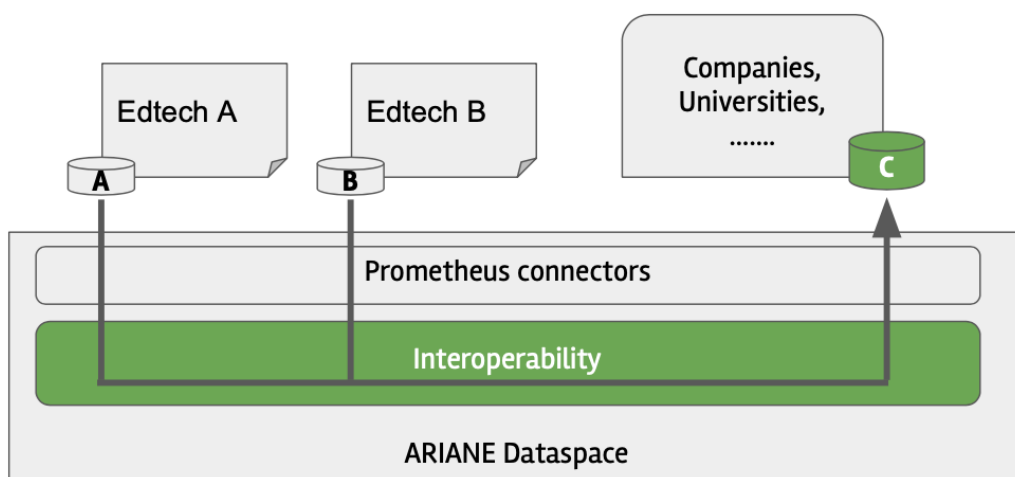
3.1 Overall software design

3.1.1 The 2 purposes of the software

The first purpose of the software makes it possible to retrieve data from different sources and to have the ability to read them in the same format (Notion of pivotal Ontology defined below) and with competence values from an authority framework .

Thus, we can have as an example, a university which wishes to recover the data of a student in terms of possible skills and orientations. To this end, we can imagine that a first Edtech had the student take personality tests and that the second offered them different types of possible orientation based on the results of this test. The university will have access to all of this information in the same format and with values from a framework that it controls (ROME for example).

This is what the Interoperability module offered by the ARIANE Dataspace services through the Visions access and consent connectors.



A-B : specific data in a specific format

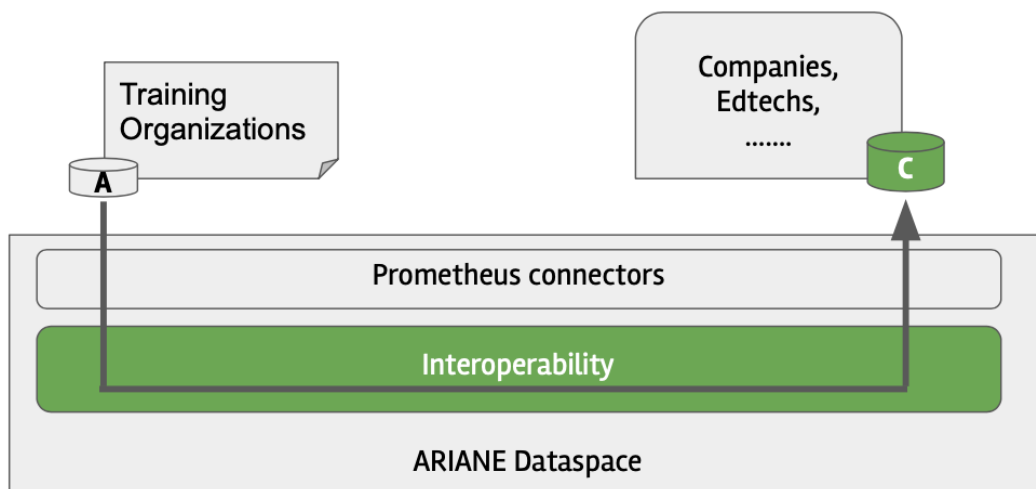
C : combination of data (A and B) in a common format : **Pivot Ontology**

C : data are natives or aligned with authority **Frameworks** (ROME, ESCO, ...)

The second purpose of the software makes it possible to retrieve training data potentially augmented with skills and skill blocks. The idea here is that training organizations have the opportunity to be part of the ARIANE Dataspace ecosystem in order to make all of their training data accessible in the same format (Notion of Pivotal Ontology defined below) and with skills values derived from an authority framework to enable the most widespread use possible.

For example, a company whose employee profile is defined with the skills of the ESCO framework would like to find training on the market to help them progress by skill blocks.

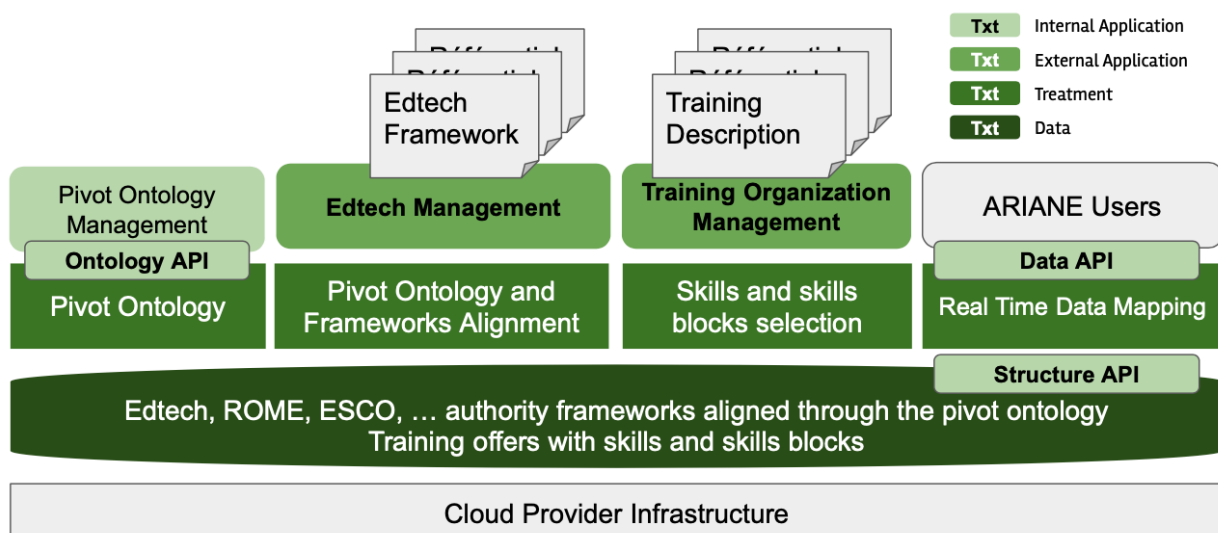
This is what the Interoperability module offered by the ARIANE Dataspace services through the Visions access and consent connectors.



The general and exhaustive specifications of the software to be produced are indicated below.

3.1.2 Software modules

The software "**Interoperability of skills frameworks**" is available in different modules as shown schematically here. The modules are described in the following paragraphs.



Internal applications are accessible only to duly authorized Prometheus administrators.

External applications are accessible to all parties authorized by Prometheus.

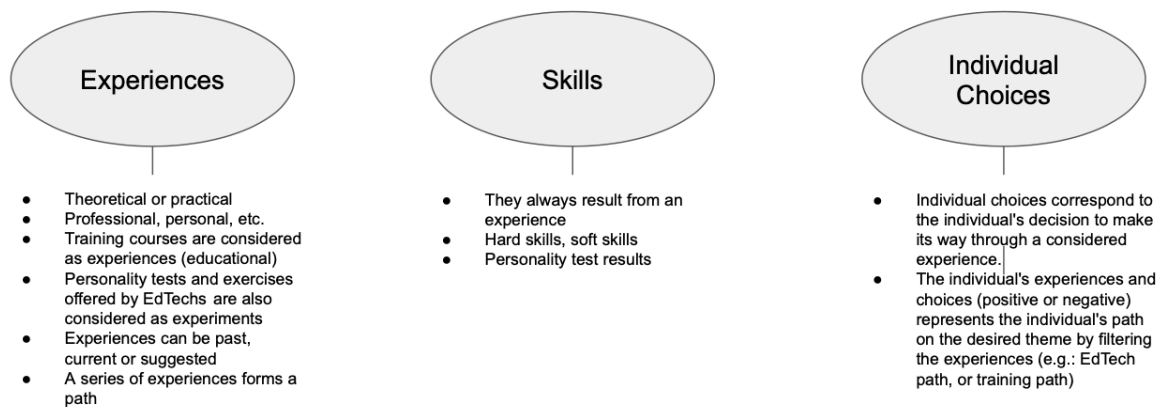
The treatments are only accessible by the developers who maintain them.

The data corresponds to the effective storage on the cloud of the various information necessary for the project (frameworks, alignments, description Edtech, description of Training Offers, , ...)

3.1.2 The Pivot Ontology and authority frameworks

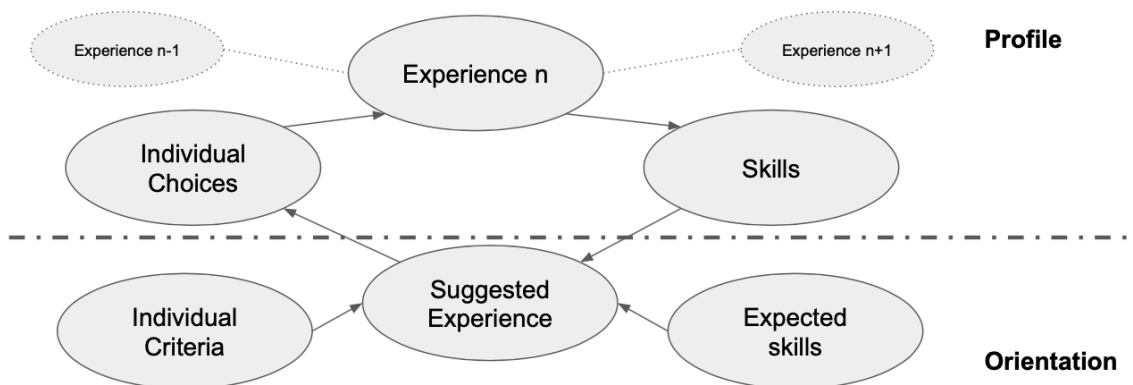
The pivotal ontology that we call “SOO v1.0” comes from Mindmatcher’s previous work within Prometheus. It describes the interactions between the main concepts capable of responding to the challenges and needs expected of Prometheus in terms of skills profile and business orientation.

The SOO v1.0 ontology is the result of an experiment in a real situation aimed at integrating and aggregating data from around ten EdTech companies focused on professional guidance through skills. The academic ontologies available in this field quickly proved to have an inadequate level of complexity and constraints given the multiplicity and originality of the data models specific to each EdTech. To respond to this problem, we designed a synthetic and compact ontology model, based on three main classes, experiences, skills and individual choices (see figure below).

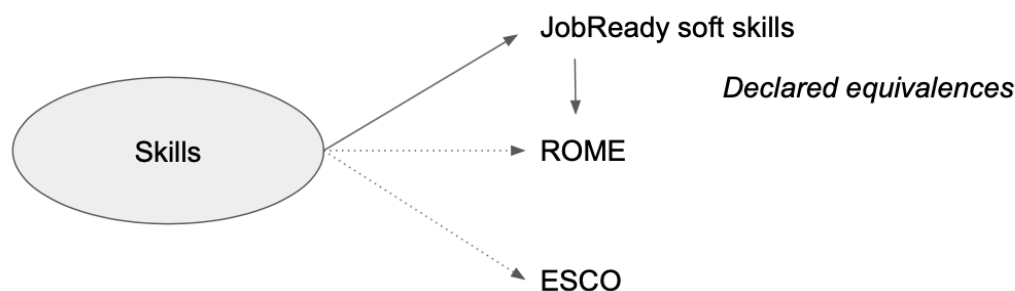


To adapt to the wealth of EdTech data, each of these classes is equipped with an original set of properties allowing it to represent all the diversity of the models encountered. For example, training is modeled in our ontology as an “experience” object whose property is to be an “educational” type experience. This shift in complexity into the properties of the objects rather than into the objects themselves results in a hierarchical, progressive structure, which allows rapid handling by a novice user.

These three objects combine to form broader concepts such as the concept of Path, Profile or Orientation (see figure below). A series of experiences forms a journey. The skills resulting from these experiences as well as the individual choices which determined them form a Profile. Finally, the experiences offered based on the match between the profile, the skills expected by the job market and the user's individual criteria (for example a geographical area) constitute our Orientation model.



Finally, to allow complete interoperability of data from the different EdTechs, it was necessary to take into account their own frameworks. Indeed, their framework, particularly of skills, constitutes a differentiating and original element of each EdTech. To preserve it while guaranteeing data interoperability, our ontology has the capacity to represent alignments between frameworks. Thus, each EdTech has the possibility of declaring an alignment between its own framework and one or more authority frameworks, such as ROME or ESCO.



This ontology will, as part of the Ariane project, evolve to take into account the challenges of this project through Workshops which will be conducted with the various interested stakeholders in the project (see below).

The pivotal ontology resulting from the ARIANE project will be called “GAIA-X Ontology”

3.1. Functional specifications of the modules to be produced

The general and exhaustive specifications of the modules to be produced are indicated below.

3.1.0 “Pivot Ontology Management” module

This module is intended for ARIANE Dataspace administrators to modify the Pivot Ontology.

A simple human-machine interface allows the pivot ontology to be viewed and modified.

Any modification of the ontology can have a significant impact on the alignments of the format of Edtec and training organizations already achieved. These modifications must therefore obtain a consensus from the ARIANE Dataspace ecosystem before being applied.

3.1.1 “Edtech Management” module

This module is the Edtech management interface with the possibility of:

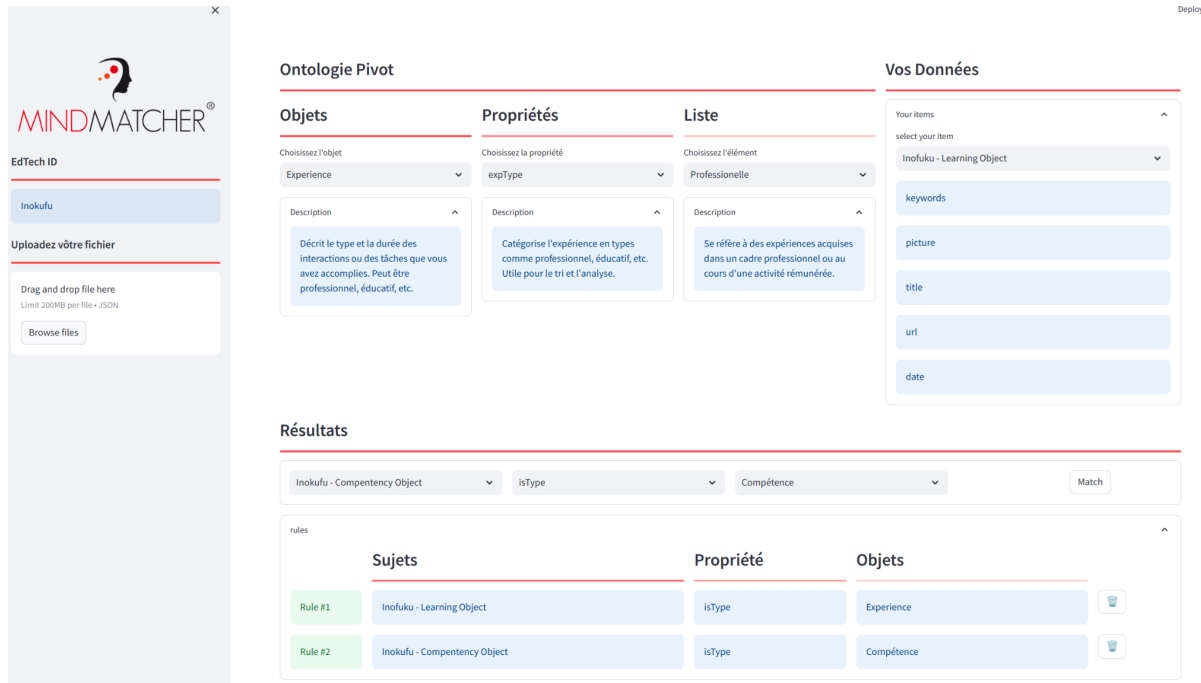
- List existing Edtech
- Create an edtech
- Delete an Edtech.

Once the EdTech has been created and to be able to use the pivotal ontology, you must be able to match the objects manipulated by the EdTech (for example the results of a personality test) with the objects of the ontology (in this case the skills associated with this test).

To do this, we will develop a HMI (see example below) allowing each EdTech to match the fields of their data model and the objects of the pivot ontology. Part of this HMI presents the user's data as retrieved

from their data file. Another part will present the different classes of the pivot ontology and the associated properties.

The workshops, planned as part of the project, will also make it possible to iteratively test and improve the ergonomics of this HMI.



The screenshot shows the MINDMATCHER web application interface. On the left is a sidebar with the MINDMATCHER logo, an 'EdTech ID' field containing 'Inokufu', and a file upload section titled 'Uploadez votre fichier' with a 'Browse files' button. The main area is divided into two panels: 'Ontologie Pivot' and 'Vos Données'. The 'Ontologie Pivot' panel has three columns: 'Objets' (with a dropdown for 'Experience'), 'Propriétés' (with a dropdown for 'expType'), and 'Liste' (with a dropdown for 'Professionnelle'). Each column has a 'Description' field. The 'Vos Données' panel has a 'Your items' section with a dropdown for 'Inofuku - Learning Object' and several input fields for 'keywords', 'picture', 'title', 'url', and 'date'. Below these panels is a 'Résultats' section with a 'rules' table. The table has columns for 'Sujets', 'Propriété', and 'Objets'. It shows two rules: 'Rule #1' with 'Inofuku - Learning Object' as the subject, 'isType' as the property, and 'Experience' as the object; and 'Rule #2' with 'Inokufu - Competency Object' as the subject, 'isType' as the property, and 'Compétence' as the object. A 'Match' button is located to the right of the 'rules' table.

The features are as follows:

- retrieving a file containing data structure in different formats
- manual alignment to the pivot ontology
- modification of alignment towards the pivotal ontology

The alignment between the data model of each EdTech and the pivotal ontology makes it possible to match a specific type of data from an EdTech (for example the soft skills resulting from a test offered by an EdTech an object of the pivot ontology, in this case a soft skills type skill associated with a personality test. It is a pairing on the container (know-how to know-how) but not on content.

Indeed, there are many possible frames of reference for knowing how to be. Another module is therefore necessary to relate the contents of the objects to each other, for example to connect the know-how coming from different EdTechs, for example to duplicate them or to align them with an authority framework such as ROME V4 or ESCO. To do this, we will create a module allowing each EdTech to match the elements of their repository with several authority frameworks. These authority frameworks will serve as a pivot between the data of the different EdTechs.

This module allows you to manage an institutional profession/skills framework or one from an Edtech company.

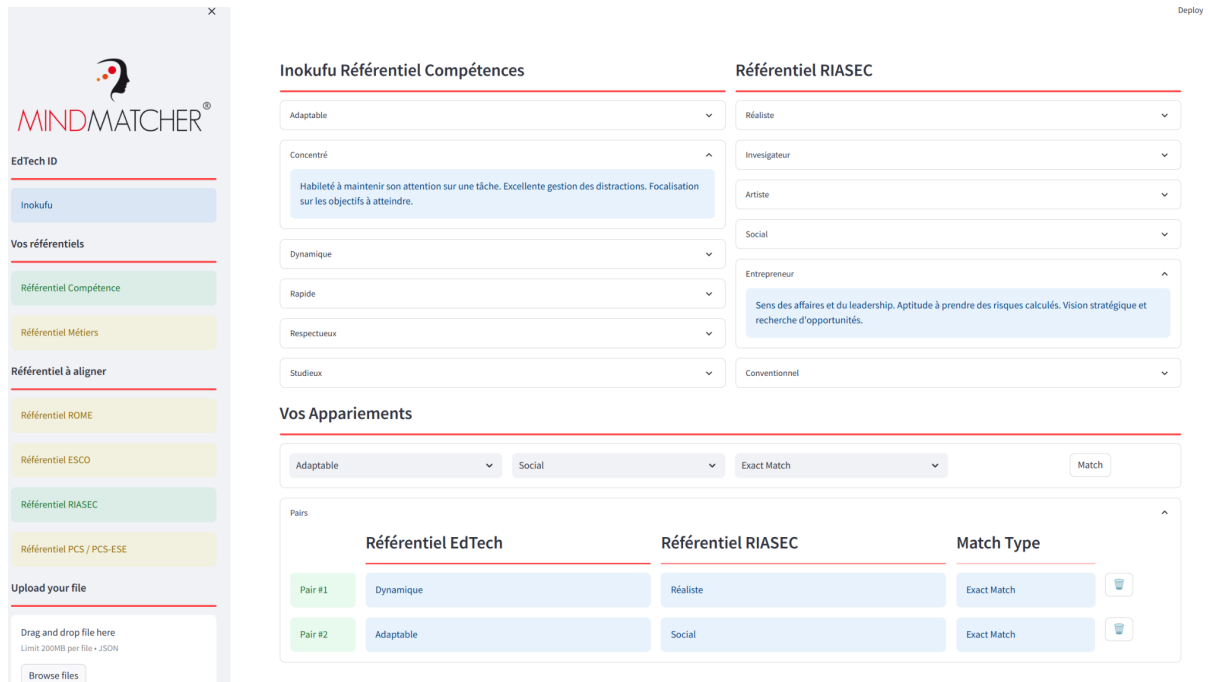
The idea is to recover the repository in question and align it with existing authority frameworks.

The features are as follows:

- recovery of a file containing the professions/skills repository in different formats



- suggestions and selection of possible alignments for each profession and each skill
- modification of the repository and/or its alignments

We will also develop an HMI (see example below) for this module. The workshops will also be used to iteratively test and improve the ergonomics of this HMI.



The screenshot displays the MINDMATCHER user interface, which is divided into several sections:

- Left Sidebar:** Contains the MINDMATCHER logo, "EdTech ID", and a list of "Vos référentiels" (Your reference frameworks) including Inokufu, Référentiel Compétence, Référentiel Métiers, Référentiel à aligner, Référentiel ROME, Référentiel ESCO, Référentiel RIASEC, and Référentiel PCS / PCS-ESE. At the bottom is an "Upload your file" section with a drag-and-drop area and a "Browse files" button.
- Main Content Area:**
 - Inokufu Référentiel Compétences:** A list of competencies with dropdown menus for selection. The "Concentré" competency is highlighted with a description: "Habilité à maintenir son attention sur une tâche. Excellente gestion des distractions. Focalisation sur les objectifs à atteindre."
 - Référentiel RIASEC:** A list of personality types with dropdown menus. The "Entrepreneur" type is highlighted with a description: "Sens des affaires et du leadership. Aptitude à prendre des risques calculés. Vision stratégique et recherche d'opportunités."
 - Vos Appariements:** A section for creating pairs. It includes dropdowns for "Adaptable", "Social", and "Exact Match", along with a "Match" button. Below this is a table showing pairs of competencies and personality types.

Pairs	Référentiel EdTech	Référentiel RIASEC	Match Type	
Pair #1	Dynamique	Réaliste	Exact Match	
Pair #2	Adaptable	Social	Exact Match	

3.1.2 “Management of Training Organizations” module

This module allows you to manage a training offer or the catalog of training offers from a training organization. As such, the main functionalities are as follows:

- recovery of a file containing the training catalog in different formats
- suggestions and selection of skills and skill blocks associated with each training
- modification of skills and skill blocks associated with each training
- deletion of the training catalog

The proposed skill blocks will be suggested as a priority with regard to the RNCP skill blocks. The man-machine interface to ensure this management will be presented during a workshop to improve its ergonomics.

3.1.3 "API Ontology" module

This API allows read-only access to all or part of the description of the pivot ontology:

This API allows you to:

- Read an extraction of all the ontology
- Read an extraction of a subset of the ontology

3.1.4 “API Structures” module

This API allows you to query the structure of authority frameworks and the various Edtech framework and training offers in a read-only manner.

We can thus:

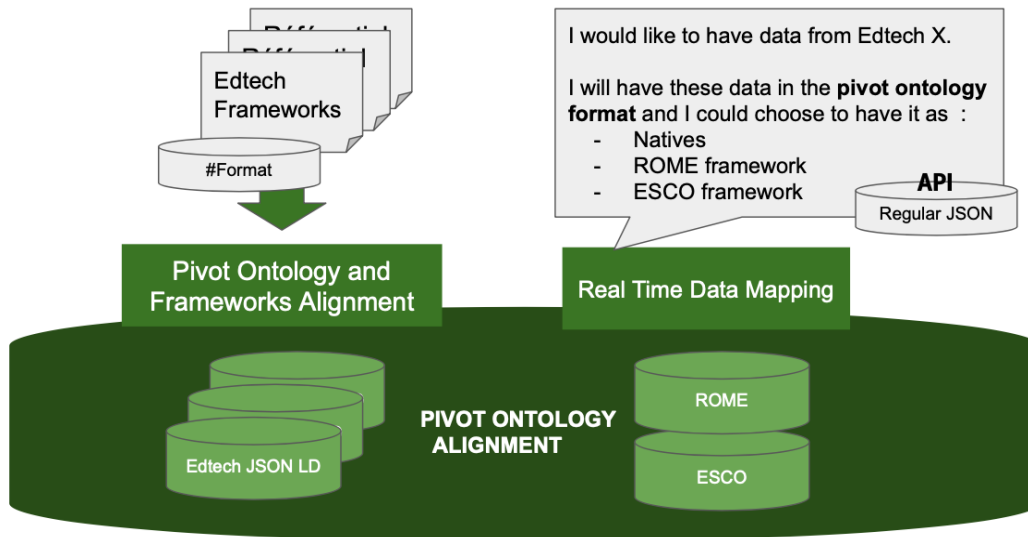
- Request the skills of a job in one of the frameworks
- Request alignment of this or that jobs or skills with authority frameworks
- Make more complex auto-completion requests across all frameworks
- Have a list of jobs with a certain number of skills
- ...

3.1.5 “Data API” module

This API ensures, in real time (Real Time Data Mapping), the translation of data from an Edtech to the pivot ontology: .

This API allows you to:

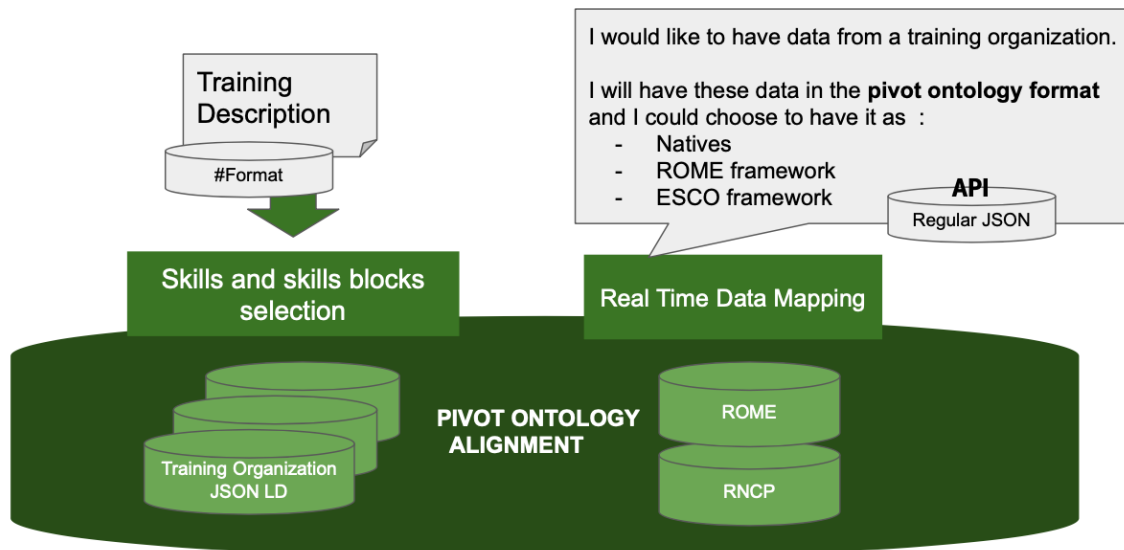
- Retrieve data from the native Edtech framework,
- Retrieve data from the ROME framework,
- Retrieve data from the ESCO framework.



This API also allows you to retrieve training in the form of queries.

So it is possible to:

- Request all training from a training organization with its native data,
- Have the same information under the ROME framework,
- Have the same information under the ESCO framework.



3.2. Ergonomic specifications

Workshops will be conducted to validate the pivot ontology and the screens of external applications

3 Workshops are planned with a maximum of 10 participants per Workshop in order to be productive over a period of 2 months (November - December)

These workshops are run by Mindmatcher and participants are invited no later than one week before the workshop. The list of participants is established by Prometheus which also ensures the invitation.

Prometheus also provides Mindmatcher with a “Miro” type collaboration tool to ensure interactive participation of people during the workshop.

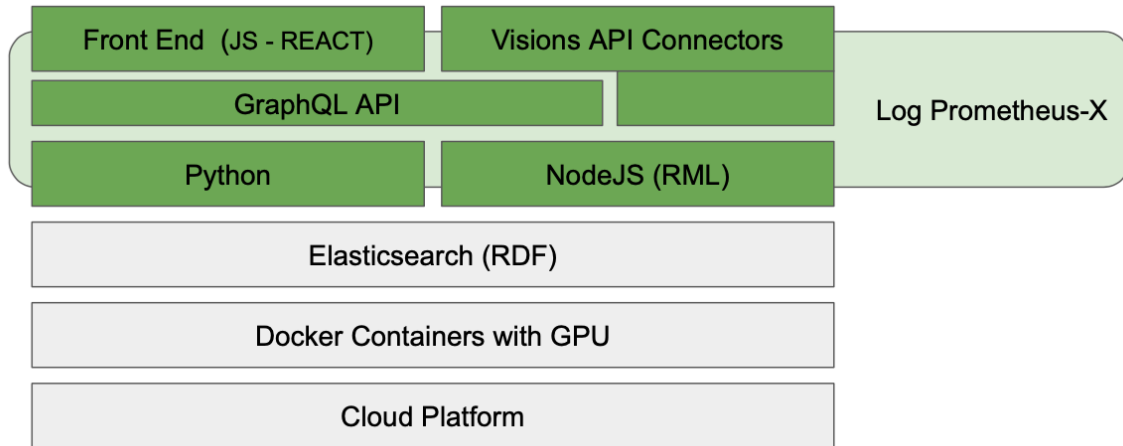
We planned to carry out 3 workshops as follows:

1. The **first Workshop** consists of a presentation of the issues, the objectives of the Interoperability software as well as a presentation of the genesis of the pivot ontology and its methods of use. Examples of dummy data are presented to understand the alignment of Edtech data with the pivot ontology and with the ROME and ESCO authority frameworks, followed by a question/answer moment and initial feedback on overall ergonomics of the “Edtech Management” module.
2. In the **second Workshop** Above all, these will be practical exercises in small groups of two or three people with practical implementation of the mapping of their own data by the participating EdTechs. It will be a question of determining which objects they manipulate in the ontology and which authority frameworks could be the subject of alignment. A final debriefing time with all participants will conclude this second Workshop.
3. The **third Workshop** allows version v0.1 of the software to be made available, namely the “Edtech Management” module. Tests will validate nominal operation taking into account feedback from the first 2 workshops. A second part of this Workshop will allow us to understand the man-machine interface of the “Training Organization Management” module with the notion of selection of skills and skill blocks.

3.3. Technical specifications

3.3.1 Main technical components

The main components of the “Skills repository interoperability” software are summarized as follows.



Front end - JS - REACT:

React is a JavaScript library developed by Facebook for creating interactive user interfaces. It allows the UI to be divided into reusable components, thus facilitating maintenance and updating. We use this technology for the development of all of our human-machine interfaces.

Visions API connectors:

These are the access connectors to the different modules of our software and in particular the Data API module. These connectors make it possible to obtain a link of trust between the requester and the data provider before using the functionalities of “skills framework interoperability”.

API GraphQL :

GraphQL is a query language for APIs and a server-side runtime environment for running these queries. Unlike REST which exposes a set of fixed endpoints for each resource, GraphQL allows clients to request exactly the data they need.

Python :

Python is a versatile programming language that is often used for web development, data analysis, machine learning, and many other applications. Its library-rich ecosystem makes it ideal for many areas including the use of Elastic search.

NodeJS - RML :

Node.js is a server-side JavaScript execution platform. RML stands for "RDF Mapping Language", a language used to express rules for transforming structured data into RDF.

Elastic Search - RDF :

Elasticsearch is a distributed search and analytics engine. In our context, it is used to search for structured data in the form of RDF (Resource Description Framework), a standard for representing information on the web.

Docker Containers including GPU :

Docker is a platform for developing, shipping, and running applications in containers. Mentioning "including GPU" indicates that some of these containers may have access to GPU resources for operations requiring significant computing power.

Log Prometheus-X :

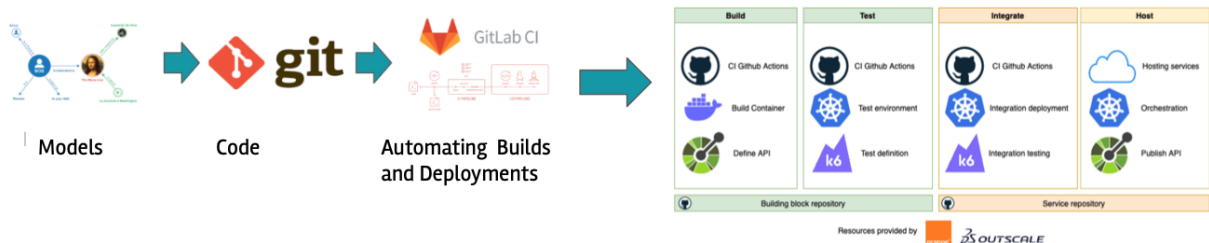
Prometheus-X is the monitoring and alert system which collects logs, as part of the ARIANE project, from the different modules for the purposes of global integration and system usage data.

Cloud Platform :

This is the cloud computing platform where the entire application is hosted. Cloud platforms provide computing resources over the Internet, providing flexibility, scale, and resilience.

3.3.2 Production chain

The following diagram shows our production line.



As part of our usual management and development practices, we use the Gitlab tool out of activism for open-source and the quality of the proposed solution.

All our developments and agile project management will pass on this platform. However, we will meet the requirements of Prometheus-x on the Github tool and the Orange deployment environment as follows:

- For code management: we will therefore use Gitlab for the day-to-day management of our developments, but we will carry out regular GIT synchronization between a stabilized branch of our Gitlab and the Github Prometheus-x repository.
- Regarding low-level artifacts (Javascript packages and Python dependencies). We will use our usual tools (Gitlab & GCP registries). We might be open to additional synchronization of these artifacts low levels on the Prometheus-x infrastructure, but on condition of technical simplicity of implementation.
- As for the final artifacts (docker-images, external dependency configuration script), these elements will be available and made available to the Prometheus-x infrastructure as appropriate. The process build of these images will be -as to do it - carried out by default on our usual infrastructure and could be transferred to the technical environment depending on the technological compatibility of the two platforms.

For our part of the process (from code to compiled docker image) the following will be provided:

- source code synced to Github
- Compiled Docker image ready to deploy in a technical environment
- Scripts for configuring external dependencies in a controlled environment

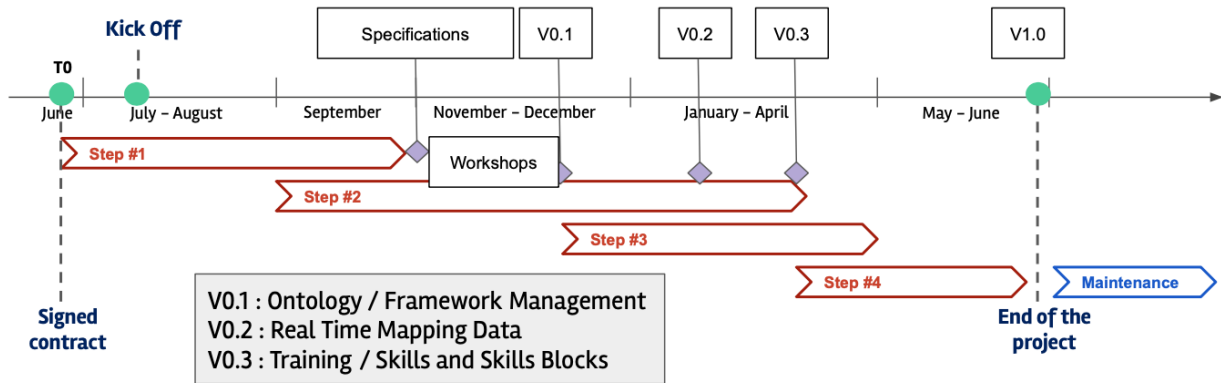
- Online development instance with its unit and/or end-to-end test reports

For the Prometheus-x part of the process: we undertake to make our best efforts to help with the technical configuration of the Orange/Outscale/Prometheus-x environment, without being responsible for possible constraints or technical difficulties of the target environment. Compliance with SLAs for pre-production and production environments is under the responsibility of Prometheus-x (mainly on the infrastructure aspect).

IV - Delivery times

The creation of the software comes in 4 steps:

- 1st step : Drafting specifications
- 2nd step : Creation and delivery of the beta version of the software (v0.1)
- 3rd step : Carrying out tests and interconnection
- 4th step : Production and delivery of the final version



Workshops (3 in number) are carried out over a period of 2 months (November - December) after validation of the specifications. The logistics and organization of the remote workshops are taken care of by Prometheus.

V0.1 corresponds to the delivery of the first version of the modules:

- Edtech management (pivot ontology mapping and alignment of frameworks)
- API Ontology
- API Structure - Frameworks

V0.2 corresponds to the delivery of another version of the modules delivered in V0.1 as well as the first version of new modules:

- Data API - Frameworks (Real Time Data Mapping)

V0.3 corresponds to the delivery of another version of the modules delivered in V0.1 and V0.2 as well as the first version of new modules:

- Pivot Ontology Management in the sense of online modification
- Management of Training Organizations
- API Structure - Trainings
- Data API - Training (Real Time Data Mapping)

At each delivery, the source of the underlying processing programs are positioned in the build space of the Prometheus infrastructure.

V - Contributions in source code

The source code contributions are as follows:

- SOO ontology v1.0 (summary of the capitalization of frameworks models)
- Frameworks Alignment AI Algorithms
- Dashemploi profile creation module (Carto.net)
- Dashemploi modules for job suggestions - offers - training
- AI algorithms for multi-support skills extraction (CV, Offers, Training)

The following links allow you to retrieve the source codes mentioned:

- <https://gitlab.com/mmorg/ismene>
- <https://gitlab.com/mmorg/>