



French Gaia-X Hub

## **POSITION PAPER**

*Data Hub « Education & Skills »*

v.1.1 – 03.05.2021

# POSITION PAPER

## *Data Hub « Education & Skills »*

<b>1</b>	<b>OUR VISION FOR A EUROPEAN EDUCATION &amp; SKILLS DATASPACE.....</b>	<b>4</b>
1.1.	Strengthening France and Europe's role in digital economy in education.....	4
1.2.	Carrying out an ambition which is politically and strategically consistent with the European vision .....	4
1.3.	Moving towards a sovereign environment.....	5
1.4.	Responding to our societies' needs.....	5
<b>2</b>	<b>TARGETED CHALLENGES 2021-2023: LIST OF USE-CASES .....</b>	<b>7</b>
2.1.	Exploring AI and data in education and guidance, with researchers: .....	7
1.	2.1.1. Modelling trends and guidance pathways to improve tailor-made counselling*.....	7
2.	2.1.2. Developing solutions to detect learners' shortcomings so as to offer personalised contingency strategies.....	7
3.	2.1.3. Developing systems to detect signals leading to school drop-out in order to offer personalised tutoring.....	7
4.	2.1.4. Promoting the use of adaptive learning in education products and services.....	7
5.	2.1.5. Developing voice assistants and learning tools for all CEFR levels and allophone students* .....	7
2.2.	Wider and better use of learning analytics by teachers and educational teams.....	7
2.3.	Strengthening the security of on-line proctoring (remote proctoring, fraud detection etc.) and developing anti-plagiarism devices.....	7
2.4.	Improving the assessment of digital solutions through harmonised key criteria description and impact measurement*.....	7
2.5.	Fostering national and international comparisons between educational systems, for example the strategies for digital skills development.....	8
2.6.	Strengthening information for families, supporting their assistance in digital use, in a school-family continuum.....	8
2.7.	Tokenisation of skills data sharing*.....	8
2.8.	« Blockchain as a service" trust platforms (in line with the European Commission's work on EBSI): trust hosting for blockchain nodes, facilitating the execution of smart contracts, legal translation of foreign diploma and diploma recognition in Europe. ....	8
2.9.	Personal education files portability for pupils, teachers and citizens.....	8



2.10. Developing a "compliance by design" framework implemented by GAIA-X governance and the Education & Skills Data Space* .....	8
2.11. Quality labelling for innovation and experimentation platforms, meeting the GAIA-X Standards, to support co-innovation and deployment of trust services.....	8
2.12. Launching a consent and ePrivacy management programme.....	8
2.13. Interoperability and/or Ontology to ensure data flow.....	9
2.14. Development of a one-stop-shop programme for data sharing for research purposes*.....	9
<b>3 ECOSYSTEM PARTNERS &amp; STAKEHOLDERS IN FRANCE .....</b>	<b>10</b>
<b>4 USE-CASES DESCRIPTION .....</b>	<b>11</b>
4.1. Use-case 1: Exploring AI and data in education and guidance, with researchers .....	11
1. Example 1: Modelling trends and guidance pathways to improve tailor-made counselling.....	11
2. Example 2: Developing voice assistants and learning tools for all CEFR levels and allophone students.....	12
4.2. Use-case 4: Improving the assessment of digital solutions through harmonised key criteria description and impact measurement. ....	14
4.3. Use Case 7: Tokenization of skills data sharing.....	15
4.4. Use-case 10: Developing a "compliance by design" framework implemented by GAIA-X governance and the Education & Skills Data Space, based on a Code of Conduct, and sovereignty and security conditions.....	17
4.5. Use-case 14 : Development of a one-stop-shop programme for data sharing for research purposes.....	19



## 1 OUR VISION FOR A EUROPEAN EDUCATION & SKILLS DATASPACE

This position paper for an "Education & Skills" Dataspace is built upon four priorities:

### 1.1. STRENGTHENING FRANCE AND EUROPE'S ROLE IN DIGITAL ECONOMY IN EDUCATION

Education is facing a digital revolution and an exponential growth in digital technology use in education, training, and lifelong learning, requiring massive digital skills development within our societies. In 2019, United States and China accounted for two-thirds of the seven billion euros of EdTech fundraising. In Europe, total EdTech investment has just exceeded 1 billion euros, mainly benefiting France, UK, and Nordic countries. Global market for learning management system (LMS) provides an enlightening example of the economic challenges involved: this market passed the USD 13.4 billion threshold in 2020 and is estimated to grow up to USD 25.7 billion by 2025, setting the scene for major developments, driven by the Covid crisis, such as the combination of distance learning processes with personalisation through artificial intelligence (AI) and the implementation of interoperability through the API'sation of services.

Digital education market is currently dominated by US, UK and Chinese players: GAIA-X initiative marks **an alternative model to data handling practices, based on a European digital vision**, and makes it possible, and within reach.

### 1.2. CARRYING OUT AN AMBITION WHICH IS POLITICALLY AND STRATEGICALLY CONSISTENT WITH THE EUROPEAN VISION

Our ambition is in line with both the European strategic position for a digital Europe and the European Action Plan 2021-2027 for digital education.

An Education and Skills Datahub will enable to:

- Strengthen EU's critical digital capabilities by focusing on key areas such as artificial intelligence" [...] and "testing and adopting trustworthy AI technologies».
- Pursue an ethical use of artificial intelligence (AI) and data in education and teacher training, by supporting research and innovation under Horizon 2020 program
- Implement Cloud-to-Edge solutions and promote sectoral data spaces, especially high-value data sets.
- Enable access to large repositories of high-quality data for public and private actors aiming to increase their productivity and competitiveness, as well as pursuing societal improvements in terms of " public services" [...] and "well-being" (for students, teachers, employees, and citizens)
- Foster cross-sector collaborations and new models to enable the exchange of digital learning services and contents, based on common standards, interoperability, accessibility, and quality assurance.



### 1.3. MOVING TOWARDS A SOVEREIGN ENVIRONMENT

---

We fully support the Gaia-X strategy which aims to impose rules leading **to a sovereign cloud environment**, based on existing solutions, but **giving data owners a full control**. That is, to:

- Depend on a European jurisdiction
- Decide on data location;
- Decide who can process data, and for what purposes.

This data owner control needs to be built upon a standardised architecture and governance mechanisms for service providers and users. Regulation on transparency, interoperability, decentralisation, and trust issues will guarantee:

- Sovereign, federated, interoperable, public or private, hybrid cloud infrastructure services, supported by current operators and secured from end to end
- **Data flow** and access to AI services within the Education & Skills Dataspace

Data Governance Act (November 2020 version) emphasises the principle of subsidiarity to facilitate data flow within cross-sectoral and EU-wide value chains, and of a harmonised legislative environment. This Act establishes a system of interoperable consent, through independent data intermediaries, strengthening data-sharing mechanisms within the different European data spaces.

We intend to seize the opportunity to **boost education data sharing in a trustworthy environment**, respectful of European regulations and national laws; we carry out the vision **of regulated EdTechs relying on an appropriate levels of data governance**. GAIA-X initiative can accelerate the development and the use of reliable data-driven services for education, study and vocational guidance, and labour market. We expect GAIA-X management to establish the broadest governance as possible, to implement a trust-based framework for cloud solution providers and to unlock potential services in education.

We believe that this global strategy, open to all market players, is the most effective way of building on existing solutions, installing trust, and accelerating the development of new services that will enable a comprehensive European offer equal, or even superior, to major non-EU players, in a sovereign, secure and ethical environment.

### 1.4. RESPONDING TO OUR SOCIETIES' NEEDS

---

An enhanced dataflow in an ethical, secure and sovereign environment will provide the right framework to develop innovative solutions, setting the stage for **major transformations of educational systems** for the benefit of **the entire educational community**:

- **For students and families:**
  - A personalised learning environment, adjusted to individual strengths and needs.



Thanks to a strong political impetus around GAIA-X in recent months - and we would like to pay tribute to the founding members, as well as the French and German governments – this initiative represents **a tangible way forward to:**

- Build, assemble, and use reliable and added-value data-based cloud services,
- Create new products/services and a new offer on the market,
- Develop a business model that complies "by design" with European regulations and values.

We consider GAIA-X as an accelerator of innovation at European scale, extending open innovation and co-construction by providing secure data sharing and artificial intelligence services at scale, in line with compliance and security issues.

**Combining the strengths of EdTechs, governments and research entities, we aim to create an education and skills data space within GAIA-X, which is going to foster Europe's competitiveness and to spread of European values improving, simultaneously, research, AI development and innovation, in a legal stable environment.**

We have described below several use cases to present some practical examples and achievable scopes of action in order to guide future choices of funding by the European Commission.

For the time being, these use-cases have been prioritized with the French digital education actors. However, we have the firm intention **of sharing them with other European members to gather several countries around this same vision and its overall goal of improving education for all.**



## **2 TARGETED CHALLENGES 2021-2023: LIST OF USE-CASES**

### **EDUCATION AND GUIDANCE**

#### **2.1. EXPLORING AI AND DATA IN EDUCATION AND GUIDANCE, WITH RESEARCHERS:**

---

- 1. 2.1.1. Modelling trends and guidance pathways to improve tailor-made counselling\*.**
  
- 2. 2.1.2. Developing solutions to detect learners' shortcomings so as to offer personalised contingency strategies**
  
- 3. 2.1.3. Developing systems to detect signals leading to school drop-out in order to offer personalised tutoring**
  
- 4. 2.1.4. Promoting the use of adaptive learning in education products and services**
  
- 5. 2.1.5. Developing voice assistants and learning tools for all CEFR levels and allophone students\***

#### **2.2. WIDER AND BETTER USE OF LEARNING ANALYTICS BY TEACHERS AND EDUCATIONAL TEAMS**

---

#### **2.3. STRENGTHENING THE SECURITY OF ON-LINE PROCTORING (REMOTE PROCTORING, FRAUD DETECTION ETC.) AND DEVELOPING ANTI-PLAGIARISM DEVICES**

---

#### **2.4. IMPROVING THE ASSESSMENT OF DIGITAL SOLUTIONS THROUGH HARMONISED KEY CRITERIA DESCRIPTION AND IMPACT MEASUREMENT\***

---



---

**2.5. FOSTERING NATIONAL AND INTERNATIONAL COMPARISONS BETWEEN EDUCATIONAL SYSTEMS, FOR EXAMPLE THE STRATEGIES FOR DIGITAL SKILLS DEVELOPMENT**

---

**2.6. STRENGTHENING INFORMATION FOR FAMILIES, SUPPORTING THEIR ASSISTANCE IN DIGITAL USE, IN A SCHOOL-FAMILY CONTINUUM**

---

**SKILLS DEVELOPMENT**

---

**2.7. TOKENISATION OF SKILLS DATA SHARING\***

---

**2.8. « BLOCKCHAIN AS A SERVICE" TRUST PLATFORMS (IN LINE WITH THE EUROPEAN COMMISSION'S WORK ON EBSI): TRUST HOSTING FOR BLOCKCHAIN NODES, FACILITATING THE EXECUTION OF SMART CONTRACTS, LEGAL TRANSLATION OF FOREIGN DIPLOMA AND DIPLOMA RECOGNITION IN EUROPE.**

---

---

**2.9. PERSONAL EDUCATION FILES PORTABILITY FOR PUPILS, TEACHERS AND CITIZENS**

---

**GOVERNANCE**

**2.10. DEVELOPING A "COMPLIANCE BY DESIGN" FRAMEWORK IMPLEMENTED BY GAIA-X GOVERNANCE AND THE EDUCATION & SKILLS DATA SPACE\***

---

**2.11. QUALITY LABELLING FOR INNOVATION AND EXPERIMENTATION PLATFORMS, MEETING THE GAIA-X STANDARDS, TO SUPPORT CO-INNOVATION AND DEPLOYMENT OF TRUST SERVICES.**

---

---

**2.12. LAUNCHING A CONSENT AND EPRIVACY MANAGEMENT PROGRAMME**

---





## TECHNICAL ISSUES

### **2.13. INTEROPERABILITY AND/OR ONTOLOGY TO ENSURE DATA FLOW**

---

### **2.14. DEVELOPMENT OF A ONE-STOP-SHOP PROGRAMME FOR DATA SHARING FOR RESEARCH PURPOSES\*.**

---

\* Six use-cases are described in section 4 below. Other use-cases' description is in progress.



### 3 ECOSYSTEM PARTNERS & STAKEHOLDERS IN FRANCE

- [Ministère de l'éducation nationale, de la jeunesse et des sports \(MENJS\)](#)
- [Académie des technologies](#)
- [Association française des industriels du numérique pour l'éducation et la formation \(AFINEF\)](#)
- [Business France](#)
- [Club informatique des grandes entreprises françaises \(CIGREF\)](#)
- [Direction générale des entreprises \(DGE\)](#)
- [EdTech France](#)
- [France Éducation international \(FEI\)](#)
- [Institut national de recherche en informatique appliquée \(INRIA\)](#)
- [Office national d'information sur les enseignements et les professions \(Onisep\)](#)
- [aNewGovernance \(aNG\)](#)

*(To be completed)*



## 4 USE-CASES DESCRIPTION

### 4.1. USE-CASE 1: EXPLORING AI AND DATA IN EDUCATION AND GUIDANCE, WITH RESEARCHERS

#### 1. Example 1: Modelling trends and guidance pathways to improve tailor-made counselling.

##### Problem

Guidance practices should be based upon efficient personalized systems: each path, in studies or professional life, is a unique individual experience. Best guidance requires using a wide data-set (local employment opportunities, personal skills and aptitudes, etc.). Most of the recommendation systems used nowadays do not exploit enough data to implement effective systems.

##### Solution

Federate guidance and skills data (student portfolios, badges, extracurricular data) to identify information on pathways, needs, skills and to build intelligent and predictive models.

##### Difficulties/barriers

- Lack of infrastructure allowing data-flow, data interoperability, legal and consent management (use-case n°9)
- Lack of standards to share skills data (use-case n°7)
- Lack of ethical framework on predictive models in guidance
- Legal Framework

##### Partners/Ecosystem/Synergies

Schools, Universities, State agency for guidance (ONISEP) and employment (Pôle Emploi), Edtech, local communities, etc. Possibility to capitalize on existing projects as Lyli project (University of Cergy-Pontoise, Académie de Versailles) or European/international projects gathering guidance stakeholders.

##### Main technology / Gaia-X

- Openness and transparency through the identification of established regulatory and data protection criteria and the guarantee of data control over use.
- Use of existing reference architectural principles, particularly those promoted by the International Data Spaces Association (IDSA).



- Interoperability at three infra-structural levels: interconnection, technical and semantic, for network, data and services.
- Facilitate intra- or inter-domain specific data exchange or data-services link beyond vendors and customers' possibilities
- Security and data sovereignty management within a framework based on GAIA-X principles certifying providers, nodes and services, particularly with regard to technical and organizational criteria. The certification will be proven by an independent and trustworthy third party and be based on already established audit and certification procedures (e.g., Minimum standards for external cloud services by BSI, C5, ISO 27001 et Trusted Cloud).

### Expected benefits

Extend guidance and recommendation possibilities throughout the European territory, and beyond.

#### **For users**

- A more precise and targeted guidance advice.
- Better visibility on possible pathways according to anyone's specific profile.

#### **For institutions and companies**

- To propose innovative, sovereign, and trustworthy offers in guidance and assistance
- To promote a know-how on international level.

## **2. Example 2: Developing voice assistants and learning tools for all CEFR levels and allophone students**

### Problem

Speech recognition quality is progressively increasing and pinpoints real “conversational” opportunities for learners, wishing to develop their speaking skills or other subjects' knowledge, like mathematics. Text recognition, errors analyses, and adaptive evaluation allow language training to be adjusted to each learner's strength and weaknesses, language competence being a global competence, measuring know-how in different competence areas, as recommended by CEFR. Similarly, voice assistants can be particularly relevant for young children reading skills. Innovative solutions exist but are struggling to develop because of insufficient language corpus to be trained upon, with a multilingual and multidisciplinary dimension, and containing a sufficient share of children's voices. Many researchers or R&D actors in automatic language processing, for instance, have major difficulties in working on most of languages, due to absence of significant data. Sometimes, although existing, data is unavailable because of data owners' difficulties in releasing (GDPR compliance, voices being personal data, children voices being even more protected).



## Solution

Use the infrastructure and framework offered by GAIA-x to promote data availability and support the constitution of common voice datasets in a learning situation (collection, reception, validation and harmonized quality processes). This implies:

- A cooperation between EdTech and data owners for data sets creation
- Open and free procedures, to make data available for research and development purposes, preventing GAFAM monopoly.
- Procedures to properly collect, process and anonymize voices
- Procedures to constitute relevant samples and to “editorialize” collected content

## Main aims pursued

### **Voice command** (NLU pour Natural Language Understanding)

- Recognition of instructions given to an application, either for control (enter a command) or for information input (answer a question)
- Cleaning the sound message to facilitate understanding in noisy environments.
- Voice transcription in applications (direct voice input into a word processor with good quality transcription, to reflect words intelligibility)

**Speech analysis** (to analyze fluency, pronunciation, grapheme-phoneme correspondence, phoneme correspondence etc.)

- Algorithm development according to CEFR levels and mother tongue
- Enrichment of the specialized APIs as SoapBoxLab
- Facilitation of API development by EdTechs

## Difficulties/barriers

- Difficulty of collecting and processing data (children’s speech data, for example) to develop effective models for learning.
- Lack of a sovereign mutualized and secure infrastructure to assemble language corpus and to ensure data flow.
- Lack of quality data: frequently, a large percentage of voice recording (up to 80%) can be unusable. At the same time, real recording conditions are essential: voice recorded in a silent environment would not allow relevant AI training for voices in a classroom, in a group or even a family context. A consistent editorial work is required post-recording to obtain proper and usable data.
- For language learning, importance to take into account mother tongue, or language of use, and to manage the diversity of accents.
- Lack of procedures to create language corpus and to deal with their governance.
  - Legal Framework
  - Contractualisation and consent management



Partners/Ecosystem/Synergies

Schools, EdTechs, language testers, universities, teachers, cultural cooperation networks, language centers, individuals

Main technology/Gaia-X

- A transparent, regulatory, and protective framework facilitating data release and control over data use.
- Interoperability at three infra-structural levels: interconnection, technical and semantic, for network, data, and services.
- Structuring governance to ensure control on data use, particularly for R&D purposes.
- Security and data sovereignty management within a framework based on GAIA-X principles.

Expected benefits

- Reference language corpus and governance procedures federating all potential data owners.
- Emergence of new adaptive and personalized language solutions offers.
- Promotion of European multilingualism and plurilingual language skills.

#### **4.2. USE-CASE 4: IMPROVING THE ASSESSMENT OF DIGITAL SOLUTIONS THROUGH HARMONISED KEY CRITERIA DESCRIPTION AND IMPACT MEASUREMENT.**

---

Problem

Digital education offer, public and private, is extremely rich and heterogenous. Digital demand, for products and services, is as specific as the offer. COVID-crisis has once again shown the urgency to capitalize on what exists, to make education systems more resilient and to preserve pedagogical continuity. Education practitioners and stakeholders (teachers, but also educational managers), and families, have real expectations of offer qualification, in terms of educational impact and targeted skills. On their side, solution publishers are struggling to gather sufficient knowledge on the impact of their products, as they have no visibility on extra-curricular continuum of each student. Any data use for a specific solution can't provide enough information to assess the impact in terms of quality of learning, without a more holistic view on on-line activity and learning "traces" (for students, for teachers etc.)

Solution

Gather data on students' learning process, collected throughout all the solutions and services used, in a school & extra-curricular continuum, to analyze the impact of each solution and identify room for improvement.

#### Difficulties/barriers

- Lack of infrastructure for dataflow and lack of interoperability
- Consensus to be found on standards to describe digital solutions for education, and skills that can be developed (use-case n°13: "interoperability, ontology")
- Legal framework (learning analytics) and consent management

#### Partners/Ecosystem/Synergies

Schools, universities, state administration and operators (CNED, CANOPÉ, FEI) Ed techs.

#### Main technology/Gaia-X

- Establishment of a regulatory and protective framework allowing data circulation and use.
- Security and sovereignty management
- Supra-national framework allowing offer standardization and, ultimately, its evolution towards better quality for users.

#### Expected benefits

- Provide the entire educational community a better knowledge on digital education offers and recommendation systems for the most suitable resources and services.
- Boosting a European offer of digital educational solutions, based on European values, guaranteeing data valorization and protection.
- Develop European EdTech offer towards interoperability and educational quality criteria, making it more competitive on the international scene.
- Empowering Europe to be more resilient, in case of crisis, and to support other countries in its international cooperation action.

### **4.3. USE CASE 7: TOKENIZATION OF SKILLS DATA SHARING**

---

#### Problem

Today, there is no platform working as a single point of entry, aggregating European citizens' skills: each training organization, employer, and higher education institution proposes its own reference system. Moreover, this data on competences is fragmented between the actors of education, training, employment and guidance without any



possibility to easily identify, circulate and recognize these competences. This leads to delays and lack of relevance of recommendations and policies.

Many innovations have aimed to equip individuals with competency portfolios based on "open badges" that allow the individual and organizations to view competences from different sources. However, these experiences have not been widely adopted for several reasons:

- **Interoperability:** A major change in the stakeholders' information systems is needed to accommodate open badges, it is therefore important to implement an automatic way to manage this interoperability.
- **Recognition:** Open badges are not governed or officially recognized; it is necessary to set governance and data traceability to ensure that each stakeholder can contextualize and have confidence in the data they receive
- **Data use:** there is no incentive to share and circulate open badges under the person's control. It is important to develop access management functionalities, nevertheless the means, the solution or the place of data storage.

### Solution

Produce different mechanisms based on artificial intelligence, usable by all stakeholders, upon skills data:

- Competence "extractors": algorithms capable of extracting competences from raw data on the basis of common reference systems (ROME/ESCO/RECTEC...);
- Inter-repository translators: algorithms capable of semantically linking competence repositories together and linking them to central repositories;
- Skills- tokenizers: based on the European blockchain (ESBI), all the mechanisms and standards for recognizing and tracking badges and skills validation.

### Barriers

- Excessive data spread: skills data are not always explicitly described and can be scattered over several storage locations and between different data owners.
- Lack of the right amount of data to make the appropriate correlations between ontologies and skills' profiles.
- Data space should provide sufficient data to train algorithms properly.

### Partners/Ecosystems/Synergies

OpenClassrooms, Simplon, Paris 1, University of Lille, Université Catholique de Lille, Hauts-de-France, Provence Alpes Côtes d'Azur, University Cergy-Pontoise, Academy of Versailles.





And internationally: Europass, Tampere university, Frankfurt Universities, Finish Defense Forces, Swedish Employment Agency.

#### Main technology/components Gaia-X

Sovereign data exchange

#### Concrete advantages/expected benefits

This type of solution will enable education, employment, training and guidance stakeholders to:

- Determine European citizen's skills;
- Be able to align their respective competency frameworks to promote matches and thus to accelerate training and return-to-work trajectories.
- Implement lifelong learning and recommendation process, based on better defined data

### **4.4. USE-CASE 10: DEVELOPING A "COMPLIANCE BY DESIGN" FRAMEWORK IMPLEMENTED BY GAIA-X GOVERNANCE AND THE EDUCATION & SKILLS DATA SPACE, BASED ON A CODE OF CONDUCT, AND SOVEREIGNTY AND SECURITY CONDITIONS.**

---

#### Problem

European education community (parents, teachers etc.) is strongly concerned about personal data protection and data use in a sovereign and ethical framework. In order to overcome any psychological or behavioral barrier and to implement a Data Space at European level, it is necessary to create a trust framework and new opportunities of products and services, based on educational data.

#### Solution

GDPR and Data Governance Act, together with the idea of a European sovereign cloud, within GAIA-X community, offer countless perspectives to reconsider place and role of data in education systems.

Our strategy is based on a sovereign and trust framework built at European level with education and lifelong learning sectors, relying on a European **"compliance by design" framework** put in place by GAIA-X governance and adjusted, by the Education and Skills Data Space members, to the digital education ecosystem needs.

This "compliance by design" framework requires to use a next generation linked data infrastructure that meets the highest requirements in terms of digital sovereignty and that



promotes innovation based on European values, such as those provided by GAIA-X initiative, related to a **Code of Conduct** of EdTechs. This Code represents a GDPR transposition to digital education sector; it is going to be proposed by EdTechs to European authorities, to give companies precise compliance guidelines and to foster a unique vision, for all European EdTech providers, on the conditions for deploying digital products and services in education.

The Data Governance Act, which is currently being finalized, is intended to enable the implementation of new rules via a common data sharing infrastructure. Some of these rules can be developed in a cross-sectoral basis, others are sector-specific. This “compliance by design” framework will allow to experiment these rules and to reinforce learners’ data security and safety.

In general, the building blocks of the future European regulation are:

- Consent management: common infrastructure and tools for managing consent on data sharing
- Identity management: mechanism linking different identities together to ensure correct identification of individuals within different services
- Governance management: infrastructure to automate compliance with governance, ethical, code of conduct and regulatory rules on the flow of education, skills, employment and training data
- Contract management: mechanism and infrastructure for generating and tracking contracts for data exchanges
- Data interoperability management: APIs and data interoperability standards
- Trust, certification and security management: certification and authentication infrastructure for organizations exchanging data
- Management of value exchanges on data: management and traceability of monetary or value exchanges linked to data exchange.

This future “compliance by design” framework will have to be followed by an implementation strategy, the adoption of sectorial standards and architectural basis to put these requirements into practice, creating products and services, in the education and skills area, and make them available for all users.

#### Difficulties/barriers

- Consensus to be found on a common Code of Conduct, and adoption by EdTechs in EU.
- Adapting the compliance by design framework to the needs of the whole digital sector, coherently with the European data governance framework (Data Governance Act)
- Clarify and promote the possibilities offered by data flow within a trust infrastructure.

#### Partners/Ecosystem/Synergies

European Ed techs, state administrations, G29, a New Governance.



### Main technology/Gaia-X

- Openness and transparency through the identification of established regulatory and data protection criteria and the guarantee of data control over use.
- Use of existing reference architectural principles, particularly those promoted by the International Data Spaces Association (IDSA).
- Interoperability at three infra-structural levels: interconnection, technical and semantic, for network, data and services.
- Facilitate intra- or inter-domain specific data exchange or data-services link beyond vendors and customers' possibilities
- Security and data sovereignty management within a framework based on GAIA-X principles certifying providers, nodes and services, particularly with regard to technical and organizational criteria. The certification will be proven by an independent and trustworthy third party, and based on already established audit and certification procedures (e.g., Minimum standards for external cloud services by BSI, C5, ISO 27001 et Trusted Cloud).

### Expected benefits

- Capitalize on work already done to develop a Code of Conduct with French EdTechs
- Transparency and common rules for all European states, based upon GAIA-X technical and conventional basis.
- Boosting a European offer of personalized digital educational solutions, based on trust and consent, guaranteeing data valorization and protection.

## **4.5. USE-CASE 14: DEVELOPMENT OF A ONE-STOP-SHOP PROGRAMME FOR DATA SHARING FOR RESEARCH PURPOSES.**

---

### Problem

Research laboratories need to access to dataset to work efficiently, ethically and transparently. Obstacles lie in the difficulty to know whether data exists or can be aggregated, to appreciate their degree of quality, to understand how to use them and under which conditions.

### Solution

The creation of a platform, gathering educational and skills data, would support most of research programs, facilitating data collect and offering the possibility to train super-algorithms (prospective, AI etc.). To summarize, the platform should:

- Collect data across the whole education and guidance ecosystem.
- Make it available to European actors (publics or private) engaged in R&D and in improving educational innovation



- Clean datasets and transform them in reference data
- Identify, define, document and share open standards to facilitate data exchange.

In our vision, this platform is not entirely public: part of the data will be open data, other part will be available for interested parties respecting appropriate entry criteria. Different access modalities can be imagined:

- Restrictive, on selected ad hoc data (like for Health Data Hub)
- Open, with a special profile for licensed professionals
- Subject to consent and agreement, as in “smart contracts systems” allowing:
  - Self-execution of a contract (fully or partially) on the European blockchain ESBI, illustrating how to link pre-determined conditions and to execute contractual dimensions
  - Data-tracking in the blockchain

#### Difficulties/barriers

- Professionals' authorization to run their services on platform's data
- Data use regulation
- Decentralized data control: large part of data is controlled by third parties independent of the MoE, especially EdTech software publishers.
- Lack of standardization (data formats differ from one publisher to another)
- Lack of interoperability: “silo effect” between solutions, preventing cross-sector data use

#### Partners/Ecosystem/Synergies

Research laboratories and universities, EdTechs, local authorities, state administrations or operators (CNED, CANOPÉ, ONISEP, FEI)

#### Main technology/Gaia-X

- Openness and transparency through the identification of established regulatory and data protection criteria and the guarantee of data control over use.
- Use of existing reference architectural principles, particularly those promoted by the International Data Spaces Association (IDSA).
- Interoperability at three infra-structural levels: interconnection, technical and semantic, for network, data and services.
- Facilitate intra- or inter-domain specific data exchange or data-services link beyond vendors and customers' possibilities
- Security and data sovereignty management within a framework based on GAIA-X principles certifying providers, nodes and services, particularly with regard to technical and organizational criteria. The certification will be proven by an independent and trustworthy third party, and based on already established audit and certification procedures (e.g., Minimum standards for external cloud services by BSI, C5, ISO 27001 et Trusted Cloud).



Expected benefits

- Datasets access to promote public and private research.
- Contribution to the establishment of a “healthy” and sovereign data economy, based on common agreement and Data Space subscription
- Digital services improvement, e.g. algorithm training to develop innovative AI solutions
- Better understanding of what is needed, to guide educational policies.
- Capacity building on the proper use of data
- Producing new data on data use, for research purposes

