Stakeholder Report on Identity Verification for Research and Education

WP5 Task2 Incubator

Stakeholder report on Identity verification for Research and Education

Contractual Date:	24/12/2020
Actual Date:	24/12/2020
Grant Agreement No.:	
Work Package	WP5
Task Item:	Task 2
Nature of Deliverable:	Other
Dissemination Level:	Open
Lead Partner:	GÉANT
Document ID:	Stakeholder Report - Identity verification
Authors:	Alan Lewis, Branko Marović, Jule Ziegler, Miika Tuisku

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 856726 (GN4-3).

Abstract
A review of the needs and issues concerning identity verification within the R&E community and the applicability of document-based solutions to address those needs and recommendations for further work.

Identity Proofing for R&E

1. Executive Summary

Identity proofing is at the centre of access to digital services and, within Research and Education (R&E), the rise in international collaboration and the increase in student mobility emphasise the related issues. Assigning a digital identity to an individual and collecting and verifying the supporting evidence as efficiently and accurately as possible are of particular significance when traditional face-to-face vetting is impractical and identity assurance is necessary.

Within R&E the task of identity-proofing falls to home organisations, although national and pan-European bodies can also fulfil this role. This study has examined the understanding of identity proofing within segments of the R&E community, the applicability of commercially available document-based solutions to address this and how the R&E community can best be enabled with this capability. Through a set of representative stakeholder interviews comprising, home institutions, research organisations, virtual organisations and NRENs/identity federations, together with secondary research we have found that, although it is early days for many in investigating the scope of this issue, there is a need for further information on the topic and a potential emerging need to support identity proofing of 'foreign' researchers and students more strongly.

National/pan-European eIDAS eIDs and federated student and researcher identities, incorporate or could serve to deliver identity-proofing and are thus often perceived as solutions in this area. But since document-based identity proofing in a large part must be handled locally, it is worth considering on its own, as well as the national or pan-European solutions that are or will be available in the short to medium term. This may assist home organisations (who are the principal providers of identity proofing within R&E) and others with addressing this problem.

The report makes four principal recommendations to further establish the scope of the issue and to help define viable solutions.

- Establish a platform for information capture and exchange on document-based identity proofing
- Create a comparative analysis of available commercial solutions
- Survey the R&E community concerning the findings so far
- Develop a preliminary business case for a community-operated service

The implementation of these is outside the scope of this report.

2. Introduction

This document describes the methodology, findings and proposals that resulted from the Trust and Identity Incubator activity to investigate how organisations within research and education are addressing the problem of identity verification. It presents the identified use cases across a set of stakeholders and attempts to determine if there is enough commonality to propose one of several approaches to dealing with the issue. These range

from providing information on the potential options, through a collaborative procurement exercise for a commercial solution, to the operation of either an identity verification service or a brokerage between several services. This work builds upon both earlier work within the Incubator activity on community-based vetting using a solution from InnoValor called ReadID, which can extract identity information from authoritative documents such as passports, and a proposal from the community that such a service provided for the benefit of research and education could be useful.

Section 3 describes the background to the potential problems and proposes several options for a potential GÉANT solution that could address the problems. For each option, a list of advantages and disadvantages are described.

Section 4 describes the potential segments that may need a solution to the identity verification problem and the use cases which apply. Also identified are the stakeholders interested in such solutions, who provided help with the study.

Section 5 details the approach used to compile the findings, which consist primarily of web-based research and interviews with interested stakeholders.

Section 6 provides a discussion of the general findings, which are grouped for convenience under several headings, and considers whether the needs could be met by one of the approaches outlined, or by other developments in identity verification, such as eIDs for individuals and specific groups, such as students.

Finally, section 7 concludes with a summary of the utility of the proposed approaches against the nature of the problem that has been understood, together with some proposals for further work.

3. Background

Context

Access to many digital services is gated by the requirement to have a given level of confidence that the identity presented is a unique representation of a known entity or individual involved in the access. Identity proofing is a fundamental stage in the sequence of events within the enrolment phase that, if successful, will allow the issuing and activation of credentials that may be used to authenticate the entity for access to the required service. Hence, identity proofing is a part of a workflow the scope of which is defined by the needs of the use case.

Identity proofing consists of three steps

1. **Capture identity** – capture and disambiguate data indicating a unique identity.
2. **Validate evidence and existence –** check documents and establish that the identity exists.
3. **Verify the match with an individual –** confirm that the identity is correctly associated with the person.

The exact nature of the evidence, verification and validation depends on the level of assurance that is necessary, which in turn depends on what is needed to access a given service or resource.

During the **capture**, identity evidence is collected to uniquely determine the identity of the subject. This evidence, such as a passport or a driver's licence, may be retrieved from one or more authoritative sources. In general, it is assumed that the issuing authority has properly ensured that the person exists and has verified the identity.

In the **validation** process, it is checked that the captured information is genuine. This can be achieved by checking the information captured against an authoritative source (for example, the issuing authority) and by checking that the captured information has not been amended. In the case of documents such as biometric passports, this can be carried out by accessing data stored securely within the embedded chip and checking the digital signature of the data.

The last step in the identity proofing operation is **verification** that the information captured and corroborated during the validation step is matched to the subject presenting it. This may be achieved remotely or face-to-face by comparing the captured information with the physical subject (if face-to-face) or a video of the subject that proves their physical existence and coherence with the information and, at the same time, confirms the person's awareness and participation. Once accomplished, the digital identity is proofed and may be used in subsequent authentications to a selected service or resource.

Within R&E, the responsibility for identity proofing typically lies with home institutions (typically a university or research community), and in general, it relies on face-to-face verification based on the information presented by the subject. Home organisations (or the NREN-operated identity federations) typically set their identity proofing policies based on their own needs, as there is currently no well-established baseline, except for the International Global Trust Fabric (IGTF) services for some disciplines at the global level.

Several standards have arisen around identity proofing and authentication assurance, such as ITU-T X.1254/ ISO/IEC DIS 29115 (Entity authentication assurance framework) and NIST 800-63A (Digital identity guidelines), as well as frameworks for assurance from organisations concerned with this area, such as REFEDS (REFEDS assurance framework) and eIDAS (eID assurance framework). Furthermore, more detailed information should be sought in these documents. Also, the GN4-3 Trust and Identity Incubator work on community-based trust establishment may also provide useful information.

Document-based identity proofing

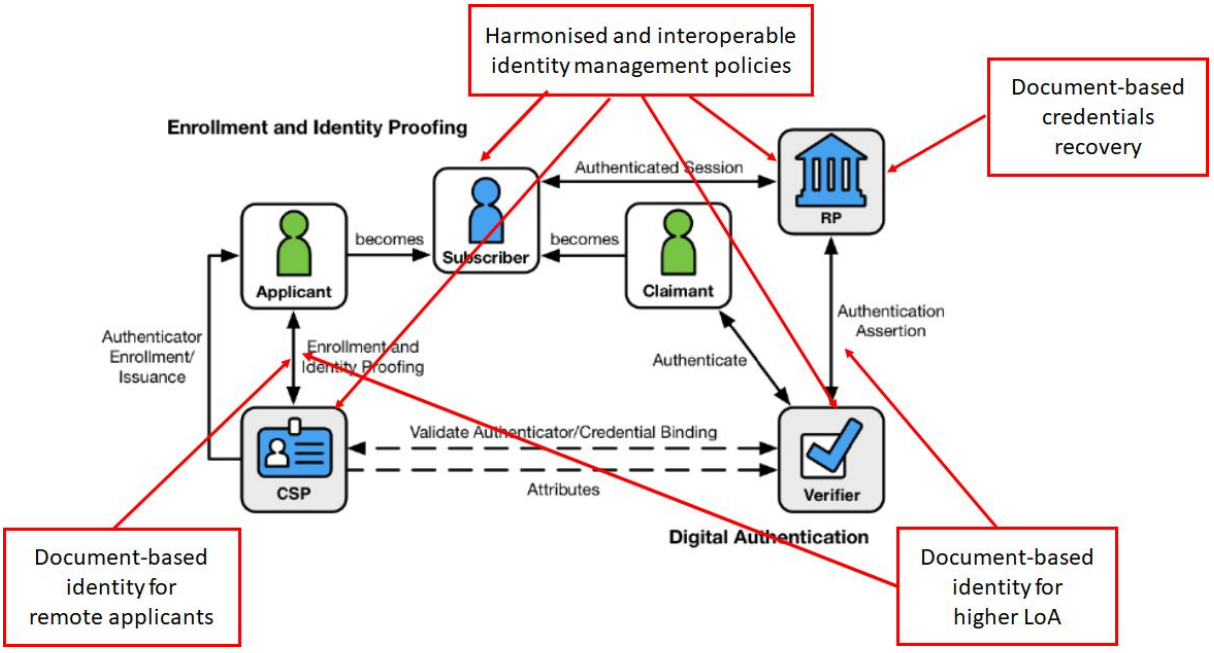
Document-based identity proofing is, in the strict sense, based on ID documents, but may also involve many types of supporting documents, such as certificates, decisions, statements, bills and letters. Many modern identity documents, such as passports and some ID cards, have an embedded RFID chip. Access to the data stored on the chip is facilitated by an optical machine-readable zone on the document. The information on the chip (e.g., the

subject's name, identifier and photo) is additionally protected by security mechanisms such as digital signatures, which also allows the authenticity of the information to be established.

Several commercial organisations offer solutions that make it possible to read documents with ICAO 9303 compliant RFID chips and often couple this with biometric technologies such as facial matching to verify that the subject presenting the document is the same as the one whose biometric information is stored in the document. This can be done, for instance, with an Android or iOS device equipped with an NFC reader and a camera. Determining the applicability of these commercial RFID document-based proofing solutions and identifying the most efficient ways to deliver them were some of the original motivations for the study described in this document.

It should also be noted that document-based identification is one amongst a number of means of proving identity to gain access to services. In particular, the increasing availability of national and pan-European (eIDAS compliant) eIDs and efforts to support student mobility, such as the European Student Card initiative, provide alternatives that may also be viable in some situations. Besides, the initiatives in self-sovereign identities (SSI) with a decentralised identity model using distributed ledger technologies such as blockchain may also play a role in the future. However, many issues are still to be resolved in this area, particularly around interoperability with existing systems.

There are several points within the enrolment and identity proofing arena in which document-based identity proofing could apply, as shown in the following diagram.



Document-based identity proofing in the context of enrolment and identity, background image source: NIST Special Publication 800-63-3

eIDAS and identity verification

Citizen identity

The eIDAS Regulation (Regulation (EU) 910/2014 on electronic identification and trust services for electronic transactions in the internal market) came into effect in September 2014 and it has been applied since July 2016. It is intended to help verify the identity of individuals and legal entities (such as businesses) on-line and the authenticity of electronic documents. eIDAS is intended to support growth within the digital arena by simplifying cross-border transactions through interoperability of eIDs across Europe and a framework for trust services such as electronic signatures, timestamps and certificates.

From September 2018, organisations offering public digital services must recognise notified eIDs from all EU members and other participating states. Currently, there are 19 notified eID schemes in 15 of the 27 EU member countries representing 58% of the population. Although the 2014 eIDAS regulation was intended to facilitate mutual recognition of trust services (electronic signatures, electronic seals, time stamps, electronic delivery services and website authentication) across the EU, it was also intended to allow access to services across the EU with an eID issued by the home country. In a recent EU report (https://ec.europa.eu/digital-single-market/en/news/egovernment-benchmark-2020-egovernment-works-people), it has been concluded that foreign eIDs currently allow access to only 9% of the services that can be accessed with a local eID. Mutual recognition of eIDs requires the deployment and testing of the national eID nodes and software. Currently, not all national eIDs nor a sufficient percentage of the population have been covered. A summary of the state of play as of September 2020 is given in a EUSMART study (https://www.eurosmart.com/implementation-of-the-eidas-nodes-state-of-play/).

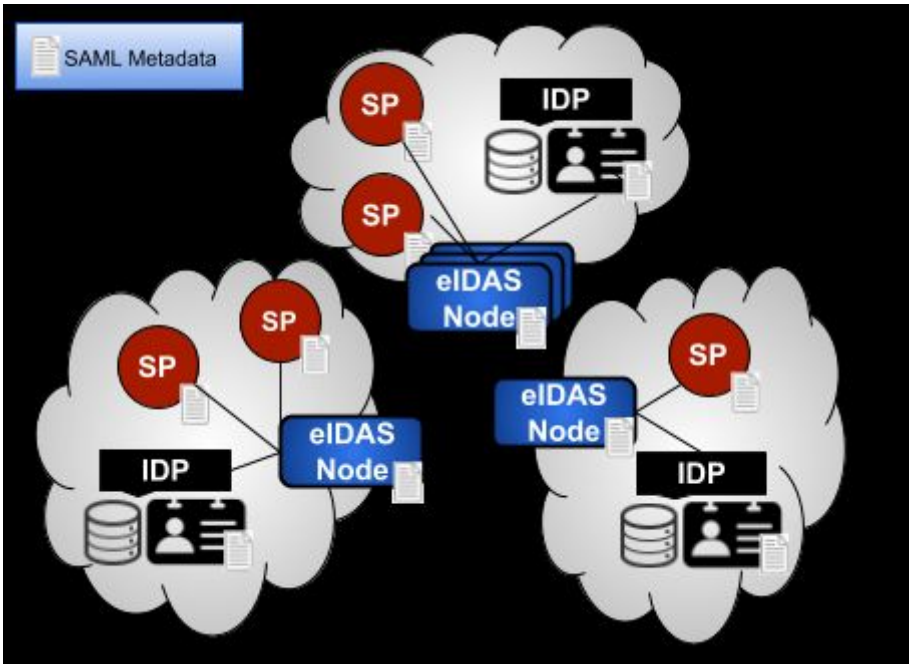
For the use cases examined, if all individuals from the EU were equipped with an eID interoperable across all EU member states, it might be considered that the need to extract identity information from physical ID documents would diminish with the adoption of eID suited solutions beyond the public services. However, it would persist for a sizable portion of foreign students and researchers.

Aside from the fact that not all member states have deployed nodes, there are also several barriers that will need to be addressed before eIDs can support the use cases. For example, the lack of persistent identifiers makes matching details received via eIDAS authentication against national records difficult. In addition, a limited set of mandatory attributes are provided. The eIDAS mandatory attribute profile only provides the following mandatory fields:

- FamilyName
- FirstName
- DateOfBirth
- PersonIdentifier

Most enrolment processes also need to check the nationality and place of birth of the individual, which is not provided but is available from documents such as passports. Finally,

the rate of growth of the eIDAS infrastructure may be slower than hoped for due to interoperability issues between different implementations and the necessary communication and harmonisation that must take place to make eIDs mutually acceptable across borders. It is possible to extend the eIDAS infrastructure to add additional academic attributes as it has been suggested by several studies [Enhancing University Services by Extending the eIDAS European Specification with Academic Attributes, https://www.mdpi.com/2071-1050/12/3/770/htm] and implemented in MyAcademicID, but the path of their adoption and population may be long and steep.



Federated Identity Architecture of the European eID System

Along with the issues concerning the deployment of functional and interoperable infrastructure, there are also concerns about the patchy nature of the uptake of eIDs across member states. In some states, e.g., Belgium, the use of eID is well established and all citizens have a unique electronic identity that is carried either on a smart card or more recently the 'itsme' mobile application to access a wide range of services. However, attempts to establish such schemes in other countries, such as GOV.UK Verify, were dramatically unsuccessful.

In an attempt to speed up the progress of digital transition, partly prompted by the COVID-19 pandemic, the EU has recently begun moves to pave the way for legislation for a secure

electronic ID system to enable all EU citizens the ability to check their identity online. Proposals are expected by mid-2021, but given the rate of progress on eIDAS over the last six years, it could still be many years until a ubiquitous eiD system is available across the whole of the EU.

Even in the event where eIDAS and its successors deliver on the promise of an interoperable EU-wide identity, the situation concerning non-EU nationals is still unclear. According to EU statistics, https://ec.europa.eu/eurostat/statistics-explained/index.php/Learning_mobility_statistics 1.3 million students from abroad were involved in tertiary level studies across the EU-27 in 2018 and a full 5% of EU residents are non-EU citizens, equating to some 22 million inhabitants. Across the world, identity verification systems range from centralised governmental systems (as in the EU and India) to private providers (such as Google, Facebook, Tencent etc. in the U.S and China) and governmental/private partnerships (as with Nordic BankIDs and the U.K.'s GOV.UK Verify). However, even governments procure 'strong' eID/eIDAS solutions from the private market. The utility of a particular system for electronic identification depends very much on the level of assurance needed for a particular transaction and views on this differ between countries. Also, cultural and political differences concerning the necessity and privacy-protective nature of an electronic ID system may mean that global interoperability is hard to achieve.

Student identity

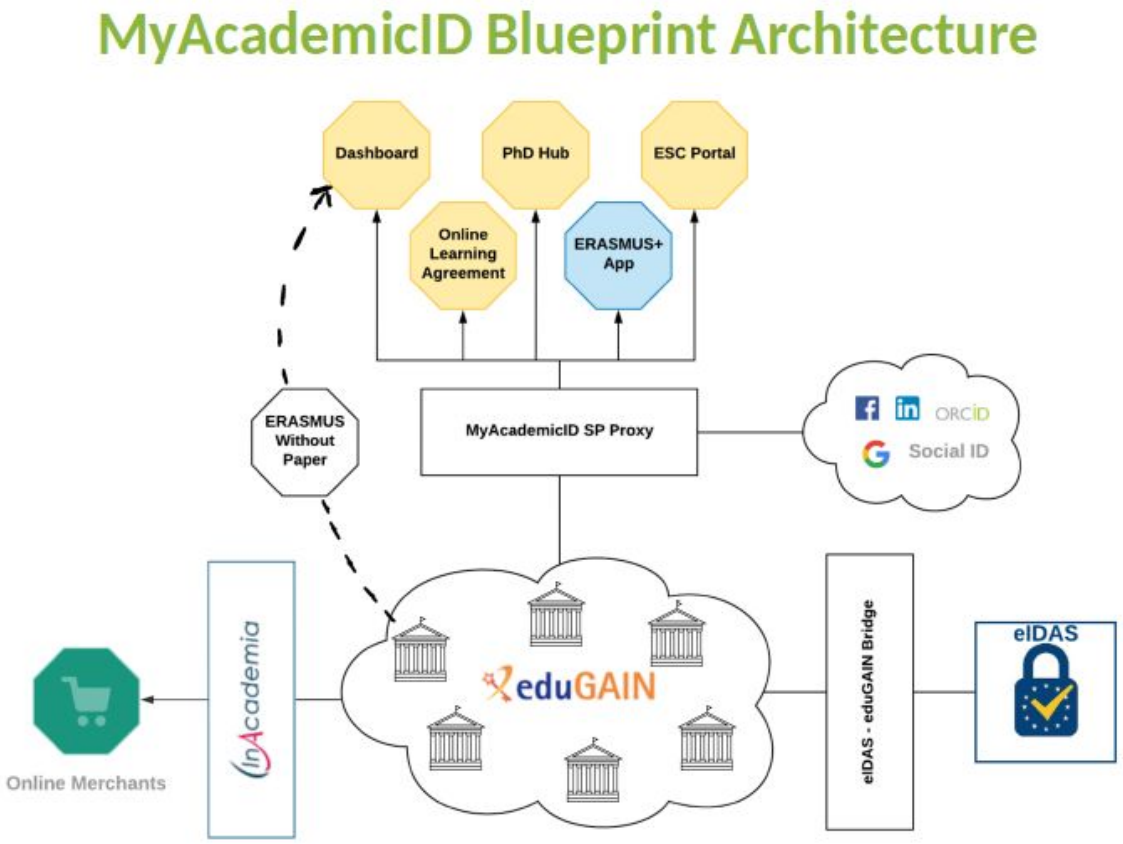
The single digital gateway enabled by Regulation (EU) 2018/1724 aims that by December 2023, at the latest, 21 important administrative procedures will be fully available on-line and national procedures will be available to users from other countries. This includes applying for admission and funding to a tertiary educational institution including requesting academic recognition and proof of study.

In parallel with this work to support the European education area (which by 2025, amongst other things, aims to make educational qualifications recognised across Europe and encourage study abroad), the European student Card initiative is intended to standardise the student mobility processes and provide a 'European student card' to allow online authentication of students across Europe. The plan is to begin deployment in 2021 and for all institutions participating in the digital Erasmus+ programme to support the full exchange of student records by 2023, with a roll-out to all students by 2025. This initiative will be linked to eIDAS online authentication.

As the roll-out of eIDAS proceeds over the next few years, it is anticipated that students will also have a national eIDAS eID that they will be able to use for enrolment in their home institution. Currently, home institutions are the authoritative source of student identity and via inter-federations like eduGAIN and eduroam are enabled with federated access to services. Enabling the use of these new eIDAS eIDs within the existing eduGAIN infrastructure is hoped to provide a single student eID. This is the objective of the MyAcademicID project, which aims to provide a standard for a unique European Student Identifier that assists in integrating eIDs from eduGAIN and eIDAS to give a single student eID. This should simplify

the enrolment process and, once eIDAS eIDs are available, take some of the load of identity verification off the home institutions. However, home institutions will still be fundamental to the release of the necessary academic attributes.

The architecture for MyAcademicID envisages a central role for eduGAIN and leverages its adoption within HE as means for federated access to services.



MyAcademicID Blueprint architecture

Although the approach will help in breaking down silos between different identity mechanisms, several issues need to be solved before such an approach can become ubiquitous. eduGAIN is central to the proposition and currently, although there are around 4,000 IdPs globally in eduGAIN, there are estimates of between 13,000 and 26,000 tertiary educational institutions worldwide. Accordingly, a substantial increase in eduGAIN affiliations would be needed. Besides, several countries already have eduID programs which would need to be fitted into this architecture and some may be using commercial third-party IAM solutions which would need to support aspects of this approach, such as the European Student Identifier to access mobility tools. Many questions remain to be addressed and some of these will be the subject of the European digital Student Infrastructure Initiative (EDSSI, which is funded for a two-year term from the end of 2020.

In addition to the points mentioned above, the projects focused on student mobility are primarily about authentication. The responsibility for identity proofing remains with governments and home institutions.

Stakeholders

Many different institutions and organisations are involved to a greater or lesser extent in the process of identity verification and may therefore have an interest in understanding the effects of a document-based verification approach and potentially deploying such an approach. Broadly, this group includes:

- Home institutions;
- Research organisations;
- Virtual organisations;
- Identity federations.

Home institutions or home organisations (e.g., universities or research institutes) run an identity management (IdM) system for their users, who may include staff, students and visitors. Examples of home institutions are universities, libraries, hospitals and research institutes. The IdM system typically enrols identities to students, employees and other persons who have an affiliation with the home institution and integrates to the institutional identity proofing processes for the users. For many users, these processes are part of the face-to-face interactions of enrolling new students and hiring employees, but there are also circumstances where in-person interaction is not in place, imposing extra complications for identity proofing. The institutional IdM system is typically run by the home institution's IT department using an on-premise deployment or, in a small number of cases (depending on their resourcing and capabilities), using a service-based offering. The IdM system is typically connected to the institution's SAML Identity Provider, which is possibly further exposed to a national identity federation and eduGAIN.

Research organisations (e.g., CERN, ESA) are collaborative organisations of diverse types and funding models that provide access to services, resources or data for research, development, publication, technology transfer and training. These may be accessed by their staff and other users whose primary affiliation is to their home institution. They may have their system for identity verification of their users, which is operated by the research organisation, or may also rely on information from home institutions. The exact approach depends on the type of user (employee, affiliate, student etc.) and the need to protect specific access to the data or resources. In general, access to sensitive information or costly resources requires a higher level of identity verification.

Virtual organisations (VO) (e.g., DARIAH, ELIXIR) are collaborations that feature users from multiple institutions, often geographically distributed, working together for a specific purpose. The creation of a VO enables the participants to share access to the data and resources they need to carry out their work. A VO may operate its own AAI and may either rely on the home institution for identity verification or carry out this role itself. Examples of VOs are research infrastructures; in its latest roadmap, ESFRI (European Strategy Forum for

Research Infrastructures) has identified over 50 research infrastructure projects and landmarks in six domains (see http://roadmap2018.esfri.eu/).

Identity federations (e.g., SWITCHaai, Haka) are groups of organisations operating service providers and identity providers that agree to interoperate under a defined set of rules and practices. The federation itself is operated by a federation operator, who is often, but not always, an NREN. Federation operators or NRENs do not provide identity verification for their IdP members, but they can agree or set standards and could provide support for such services if a need arose. Most established identity federations are members of the inter-federation service, eduGAIN.

To capture the overall needs, the study engaged with a set of stakeholders from each of the principal groups described above:

- Home institutions
- Research organisations
- Virtual organisations
- NRENs/Identity federations

Problem statement

We suppose that an increased international collaboration within and between research communities and the rise in student mobility, coupled with the need to onboard users to the services they need through digital channels, are driving the need for a means to verify digital identities and simplify user onboarding. It is not always possible, and is also costly, for a user to attend a service desk in person to have their identity verified. Furthermore, knowledge-based onboarding (where the user must answer questions related to information held by the verifying party) is subject to defeat if that the user's PII is compromised (identity theft) and can be a barrier to the onboarding process if the users often fail to remember the information asked. Therefore, various technical solutions that can enhance, support or decouple the majority of identity proofing from physical presence are of ever-increasing interest.

Document-based verification allows for remote user identification by capturing information from an authoritative, officially recognised and valid identity document (such as a passport or driver's licence) using optical or electronic means (camera image, OCR, NFC), and then assessing its validity by recognising tampering or modification, followed by a comparison between the photo of the user in the document and a video or still image taken by the user. The use of a personal mobile handset (which is ubiquitous), as a device that allows the scanning of identity documents, coupled with improvements in biometrics such as facial recognition, enables remote identity proofing, supporting different classes of remote users. In this context, identity proofing consists of activities to **capture** information from the user, **validate** that the information is genuine and **verify** that the information is matched with a real person, so as to create a verified digital identity. Such a validated identity can then be asserted during an authentication process in order to access the required service.

Within a range of environments, there are many use cases where such an onboarding capability is needed, such as within retail, e-government and financial services, where anti-money laundering (AML) compliance and know your customer (KYC) policies need to be strongly enforced. As a result, a number of providers of identity platforms have emerged, providing solutions for identity capture, authentication, digital signatures etc.

Most R&E institutions have their own IdM/AAI system but there are several use cases centred on identity proofing. Our assumptions, borne out by stakeholder interviews, have identified:

- Onboarding of researchers into research communities, groups or projects;
  For researchers who are remote from their research organisation or are part of a virtual organisation identity proofing is a challenge. This is particularly the case for researchers, such as citizen scientists who do not have their home organisation.
- Enrolment of 'foreign' students onto a campus;
  For students from outside the country where the university is based and who cannot attend in-person the enrolment process identity proofing can be difficult and may be prone to error.
- Enrolment of students into distance learning programs;
  As more institutions make courses available online, where students never physically attend an institution, accurate remote identity verification becomes essential.
- Secure account recovery following token loss or password reset;
  In the event where a secure authentication token is lost or a password needs resetting, a secure identity proofing process is needed to establish entitlement. This becomes even more difficult if the recipient is remote.

Existing solutions to address these use cases require individuals to be physically present so that their identity can be checked by the registration authority, which may be impractical, as well as costly and time-consuming. In addition, such an approach is extremely inefficient where many individuals are involved. Other challenges stem from the lack of suitably trained staff competent to perform identity proofing and the use of ad-hoc solutions combining emailing of documents and video interviews for remote individuals, which may be slow and error-prone.

Assuming one wants to resolve the above issues, an additional challenge is that, although many (commercial) identity proofing solutions exist, the best mechanism for making this capability available to the R&E community is unclear and needs investigation to determine the most viable and appropriate business case. In general:

- The cost associated with service licensing for a single institution or research community with a limited number of users may be too high;
- The interfaces (APIs) provided may not be compatible with those used within R&E AAI systems;
- A single given solution may not support the document types that are needed;

- Institutions and communities may require additional attributes for identity proofing that are not supported by a single system;
- The terms of use may not be acceptable to the community of institutions;
- The level of assurance provided may not be suitable for the use cases considered;
- Many of the solutions are regional (e.g., eIDAS) but research collaborations are often global.

To address this problem, several different approaches are suggested here.

4. Solution options

To address the problem described above, several approaches have been investigated. These approaches are presented below in the order of increasing GÉANT involvement.

Option 1: GÉANT facilitated information exchange

Description

GÉANT would act as a central point to collect and disseminate information, experiences and even code for integration on various commercial identity proofing solutions available on the market and any that are in use or have been investigated by NRENs or their customer organisations.

This could materialise, for instance, as

- A wiki page describing and linking to the commercial alternatives and their adopters in the NREN community.
- Presentations in GÉANT events or webinar series.
- Sharing code that some organisations have found useful in their integration to the commercial provider.

Advantages

- Significant benefits for early adopters and developers with little effort.
- Information exchange alone may increase the awareness of ID proofing solutions in the NREN community.

Disadvantages

- Protracted and limited benefit for NRENs and their organisations waiting to adopt matured services and potentially challenging to assess solutions at a GÉANT level.

Option 2: GÉANT-negotiated purchasing arrangement

Description

A GÉANT negotiated framework purchasing agreement that would allow participating NRENs and their institutions to access the services of one or more commercial identity proofing providers at a substantial discount against the standard licensing and transaction rates. GÉANT would not have any technical activity or, at most, it could deliver some supporting tools/code (see option 1 above) that the NRENs or their customers can deploy to integrate with the commercial service. This approach would be similar to GÉANT's role in the OCRE (https://www.ocre-project.eu/) project.

Advantages

- May provide economies of scale benefits for NRENs and VOs needing such a capability, which is currently impossible to achieve by individual institutions or VOs.
- Could reduce cost as it removes the need for separate procurement arrangements to be entered into by each NREN and makes it possible to cover multiple identity proofing providers using a single framework agreement. May provide arrangements with more than one provider to cover the broadest set of requirements.
- A single flat fee might cover all licensing and transaction costs.
- Could help harmonize technical and operational practices from vendors, as well as at implementations.
- No technical development required by GÉANT.

Disadvantages and risks

- Potentially lengthy commercial negotiations with one or more commercial providers.
- The number of R&E participants (NRENs and VOs) may be insufficient to provide a significant discount.
- Payment 'upfront' to commercial providers may be required by GÉANT and there is a risk that the actual service uptake will not cover the costs.
- International VAT challenges and procurement for non-EU countries where GÉANT does not have registration or remit may make it difficult to apply this for many VOs, although within the EU the MOSS scheme simplifies matters – https://europa.eu/youreurope/business/taxation/vat/vat-digital-services-moss-scheme/index_en.htm

Option 3: GÉANT identity broker service

Description

A GÉANT identity validation broker service connects multiple relying parties with different identity verification service providers. It provides a technical connection between the relying party and the identity provider service and insulates the relying party from the need to

implement and update the integration with the (possibly changing) identity verification service provider. It is assumed that a **contract is in place between each relying party and the used identity verification service provider** and that it will cover pricing, service levels, liabilities, etc. GÉANT itself is not part of the contract between the vendor and the relying party, but it provides technical integration for the benefit of all.

This effort could be seen as a technical implementation layered on top of the previous scenario, where GÉANT supports the procurement of identity validation services.

Advantages

- Presents a single entry point for selecting identity verification service providers for the whole R&E community which is tailored to the individual institution's needs.
- Removes complexity for the relying parties by having a single uniform point of integration, whilst providing multiple identity verification providers.
- Identity information released can be tailored to what is required by the relying party.
- Could facilitate standardization and harmonization in usage across multiple vendors and multiple consumers

Disadvantages and risks

- May need to integrate with several identity verification service providers to cover all requirements.
- Relying parties must put individual agreements in place with identity verification service providers – so, little scope for economies of scale.
- APIs vary among identity verification providers, due to which a single uniform API may not be possible and keeping these integrations up to date might be costly

Option 4: GÉANT identity proofing service

Description

A GÉANT identity proofing service is a platform owned and operated by GÉANT, built on top of an existing commercial solution, or several complementary external service providers. The service provides a uniform interface to existing R&E AAI systems and could provide workflow possibilities to add additional assertions that may be needed for a given LoA. The service has its terms of service towards the relying parties and the underlying commercial solution (including changes to it) does not necessarily need to be visible to them. This model has similarities to the current TCS solution.

Advantages

- By linking many potential customers to the commercial identity verification provider(s) it could deliver affordability through economies of scale.
- Allows for more complex workflows than those possible with a single commercial identity verification solution, as exemplified by further assertions required in some of the implemented scenarios (such as collaborative and peers' confirmations).

- Presents a uniform interface to the relying party consistent with what is needed for R&E.
- Avoids the need to change the interface if the commercial identity verification solution changes.
- The interface can be customised for the needs of the R&E community and more easily integrated with other GÉANT services (for instance, the GÉANT identity proofing service could appear as a SAML IdP in eduGAIN).
- Relying parties do not need to enter into individual agreements with identity verification service providers.

Disadvantages and risks

- GÉANT needs to own and operate the service and so revenues must be sufficient to sustain it.
- Unclear if a single commercial identity verification service could support all the requirements.
- Supporting and integrating several solutions may be a significant burden.
- Establishing a service with a single commercial solution is likely to tilt the service implementation and interface towards its platform and hamper potential changes.

Option 5: GÉANT identity proofing and MFA service

Description

This is an extension to option 4: assuming option 4 requires an end-user to install a dedicated smartphone app for passport and face photo, the app could be later used as an MFA (Multi-factor authentication) token for the user. That would extend the service from identity proofing (for first-time users) to their multi-factor authentication (for the returning users), two different things combined as a single service.

From an end-user perspective, when they use the app for the first time, it asks them to take a photo of their passport (or read the biometric passport over NFC) and a still image or live video of their face for enrolment to the service. When they use the app again, it can perform multi-factor authentication.

From a Relying Party perspective, the RP just needs to send a SAML or OIDC authentication request to the GÉANT identity proofing and MFA service and wait for a response. If the user has already enrolled in the service, it will return their authenticated identifier (and other attributes) quickly. If the user is new to the service, the same response will be returned but this takes some time as the user needs to enrol in the service first. This functionality is therefore an extension to option 4.

Advantages

- Not much extra implementation work over Option 4 if Option 4 assumes that users install a smartphone app anyway.

- Commercial APIs need to be invoked only for identity proofing, enabling providing MFA authentication ‘for free’, i.e., no need for transaction-based service pricing for the MFA service component.
- Solves the MFA token delivery, which is a pain point for MFA; an MFA token is generated automatically during the identity proofing phase and can be used for subsequent authentication.
- May help to solve the lost credential problem, which is another pain point for MFA; if a user, e.g., loses their smartphone, they can restore the situation by enrolling again with their passport.
- The responsibility for the MFA part of the service and the corresponding mobile app could be delegated to the commercial provider.

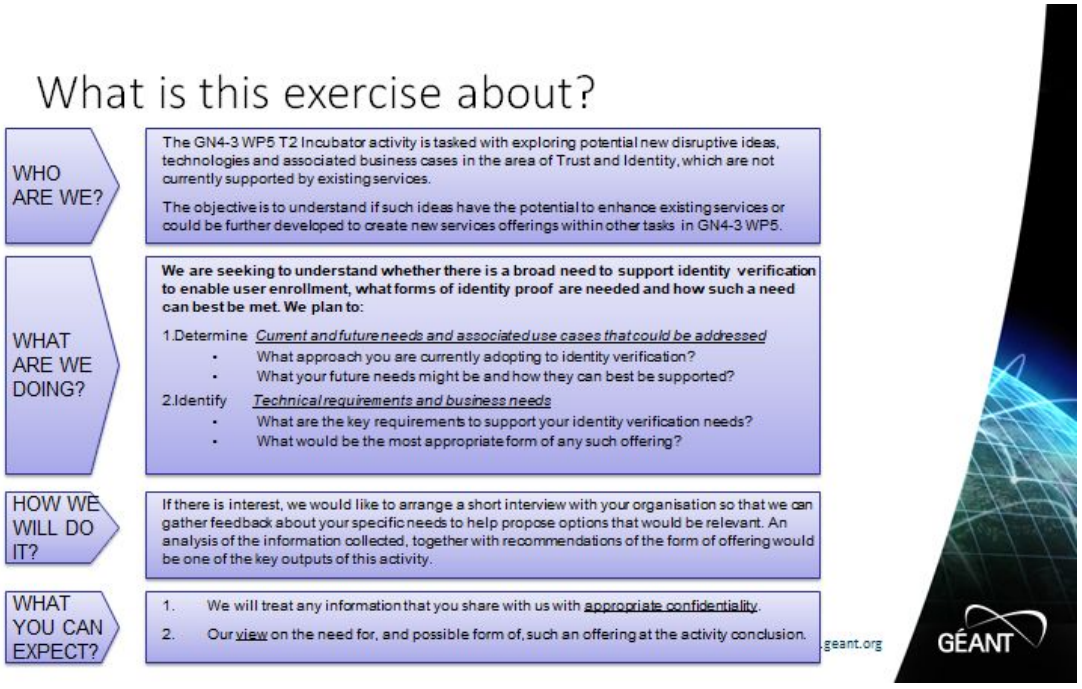
Disadvantages and risks

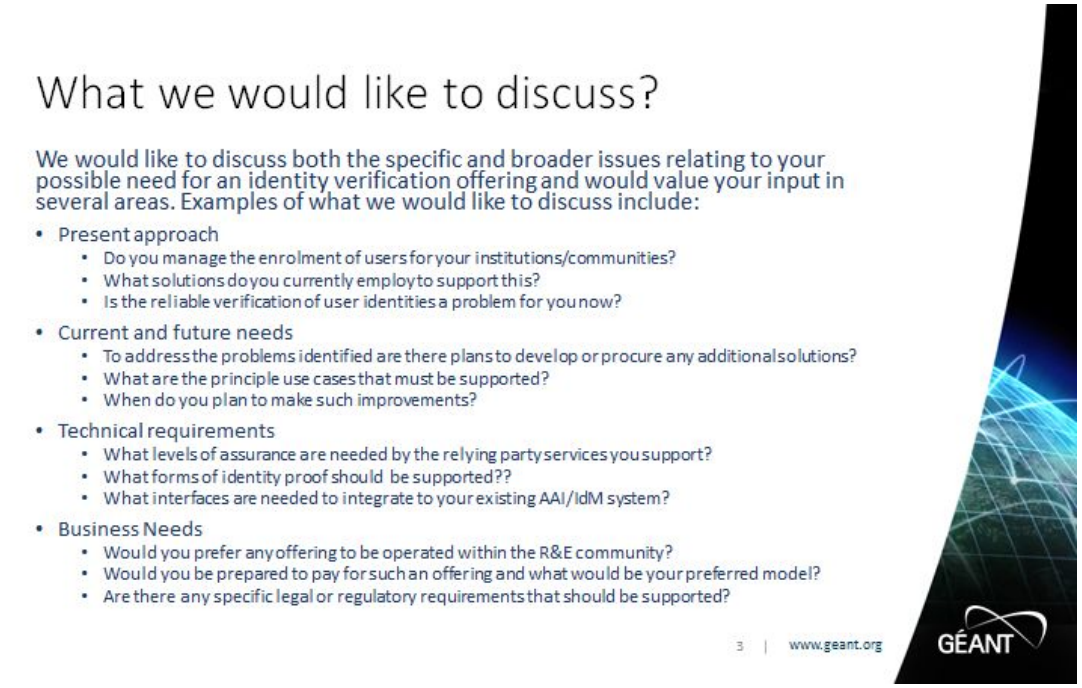
- Project scope creep.
- Increases the expectation of a long-term GÉANT service – ID validation is done just once for a user but multi-factor authentication of a returning user is performed several times.
- Entanglement with an MFA application of the commercial provider.

5. Methodology

Approach

In order to identify the needs and level of interest in the topic across the different classes of stakeholders, a series of approximately one-hour interviews were conducted with representatives of the stakeholders. Before the interviews, the stakeholders were sent some material describing the purpose of the activity and the types of information that would be covered. This helped them to consider some of the topic areas before the interviews took place.





The interviews took place over two months and during this period the questions were adapted several times, based on the experience of previous interviews. Minor changes were also made to tailor the interviews to the stakeholder involved, and not all questions were covered in each interview.

Following the interviews, an analysis took place which grouped the information under several category headings for convenience and these summaries were then shared with the respective participants to enable them to comment on or correct the information. The principal categories were:

- Scope and use cases
- Current process
- Change drivers
- Current status
- Solution landscape
- Requirements
- Implementation approach
- Business factors

Survey questions

To assist with focusing the interviews on the required areas of information, a set of questions were created to help guide the discussion. This did not preclude a freeform discussion on areas that emerged but acted as an aide-memoire to ensure the full coverage of relevant areas in the limited interview time available.

Details of the questions used in the survey can be found in Appendix B – Survey questions

6. Findings

This section presents an analysis of the interview findings grouped in different areas and attempts to highlight generalities that could indicate a common set of needs.

Scope and use cases

At the outset of the study, it was assumed that the principal use cases for document-based identity proofing would be:

1. Enrolment of (often remote) researchers into a research organisation or virtual organisation with improved LoA (e.g., beyond federated identities);
2. Enrolment of ‘foreign’ students into a university;

During the interviews, these were confirmed, as these were the most frequently cited use cases, however, there was also mention of:

3. Enhanced vetting for issuing or recovery of second-factor authentication tokens.
4. Identity vetting for password recovery
5. Enrolment of remote or short-term institutional staff and contract employees

It is not clear yet how widespread these last use cases are, as they were only mentioned by a small number of the interviewees. However, all listed cases share similar needs and elements of identity proofing. Document-based identity verification is a fundamental step in filtering applicants during enrolment and it results in a reduced need for other forms of review and a reduction in invalid applications.

Other possible use cases lie in the realm of on-line or remote learning. For example, the means of stronger identity verification may ensure that only those enrolled can attend on-line sessions, or be part of identity verification for on-line proctoring or invigilation of exams. This

last use case represents a large and growing market, but in many cases, the need is satisfied by third-party organisations that provide such services. See for example https://www.bioid.com/online-exams-e-learning/

Although several document-based identity verification solutions provide a means to use the information retrieved to enable subsequent authentications, authentication was seen as a separate step and interviewees already had systems in place to support this.

Enrolment of 'foreign' students; enrolment of remote researchers

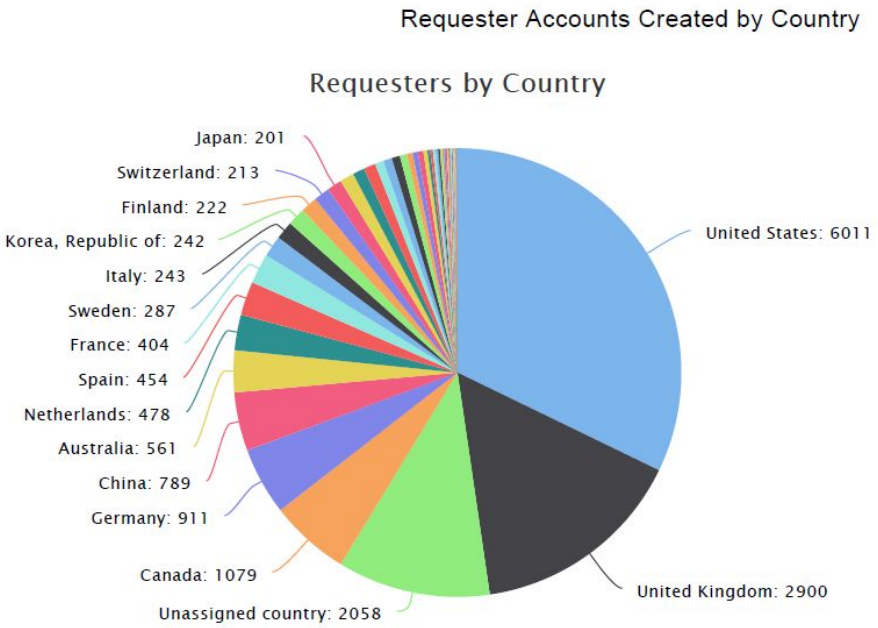
In general, the identity verification process is owned by the individual's home institution, but in some countries, there is a more centralised (national) identity mechanism. There is a range of workflows in play that can depend on the type of individual (researcher, student, employee, national or foreign) and the level of assurance that is needed. Often, enrolment is a two-part process split over time, with a less secure pre-enrolment and a more secure enrolment phase. In the case of students, this approach has been seen to give rise to some instances of fraud, where students gain an email address via pre-enrolment and then use this to prove they are students to gain student-related benefits such as vouchers from commercial organisations before they are fully enrolled. The extent to which this is practised is unknown as is the impact on institutions.

Identity verification by an individual's home institution will often involve the presentation of documentation and a F2F interview at the institution (typically the IT helpdesk to set up accounts), but depending on the regulatory environment, this may also be carried out at other locations by suitably accredited individuals. In some countries, eiDs may be used, and where they can accept eIDs from other countries (potentially by eIDAS) this may also be acceptable, although the level of assurance may be restricted. Remote vetting using techniques such as video chat for liveness verification and presentation of documents is also possible. Even in the case of an automatic enrolment process, there is a need to provide a help desk in the case of or system failure and unusual situations that may require a F2F interview.

A wide variety of documents are used as a part of the identity proofing process, including machine-readable ICAO 9303 and ISO 18013 documents issued by government agencies such as passports and driving licences, together with a range of other paper-based documents such as utility bills, tax bills, social security number cards and course certificates.

All the use cases for students and researchers could be generalised to cover all students and all researchers but the need is most keenly felt in the sub-segments of 'foreign' students and researchers from outside the EU, where other identity proofing solutions such as eIDAS would not apply.

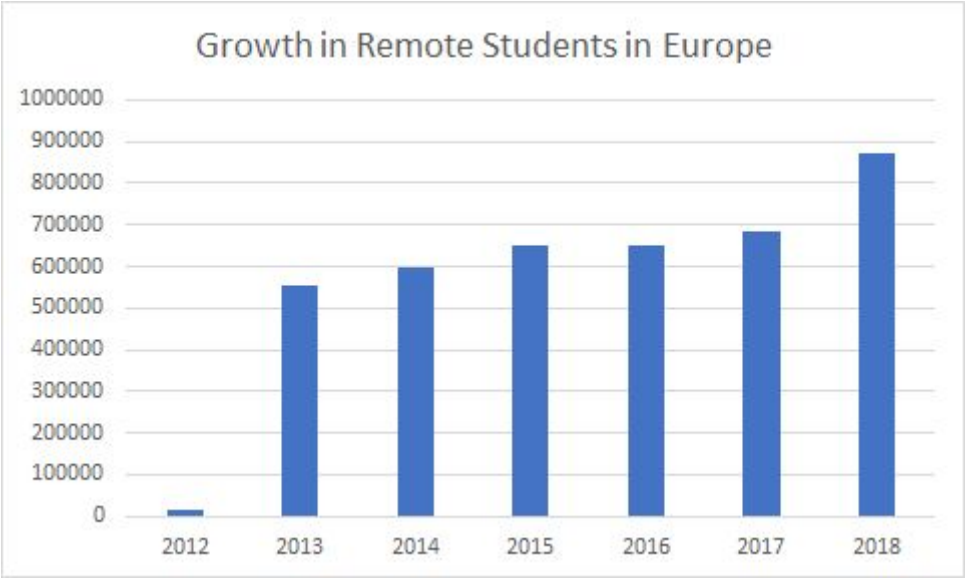
To gain an insight into the scale of these sub-segments, we examined registration statistics for several research projects.



Data from European Genome Archive

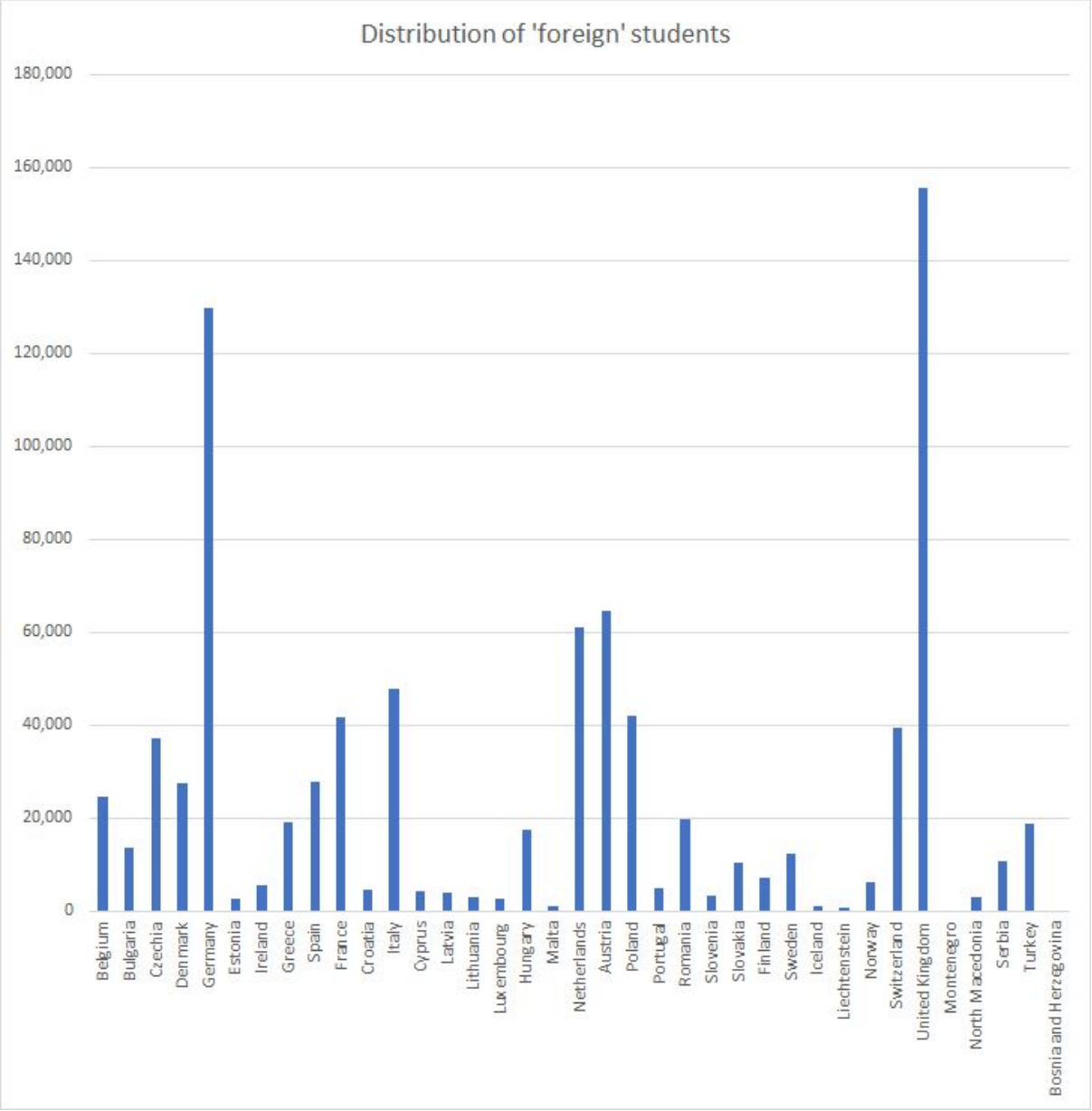
For example, data from the European Genome Archive shows that around 50% of requests are coming from outside the EU27 area where EU identification schemes are not operable.

The number of 'foreign' students who are registered for study across Europe shows steady growth, with almost 900,000 registered in tertiary education in the latest available data.



Remote students in Europe, source: EuroStat

The prevalence of 'foreign' students is particularly high in the U.K, Germany, Austria and the Netherlands.



‘Foreign’ students in Europe, source: EuroStat

This suggests that the sub-segments of researchers and students that could benefit from document-based identity vetting could be quite substantial.

Enhanced vetting for issuing or recovery of second-factor authentication tokens and password recovery

These use cases were mentioned by only a small number of participants in the interviews. A typical scenario is that the initial verification is done F2F at a service desk, but in case there is a need to recover the password, some individuals may then be remote and attending the service desk again is not feasible. In this case, the individual may be asked to upload a copy of their identity document, which is then vetted by the service desk. An automated identity

verification solution could alleviate the service desk from this task and potentially improve the quality of vetting, which would also be the case in the general enrolment use case. This requires further study.

Enrolment of institutional staff and contract employees

The use of document-based verification for institutional staff also requires more study. Given that proof of identity will only be one of the attributes necessary for an employee to enrol and it is unlikely that all the attributes could be dealt with in an automated way, it is not clear what benefit automating one part of the vetting process would bring. There could be a sub-segment of staff – for example, remote contract employees – for whom this has merit, but would be just a part of the solution providing full information.

Change drivers

There were several reasons cited why improvements in the process for identity verification could be advantageous, either now or in the future. However, for NRENs, in general, the motivation needs to be supplied by the member institutions and currently, there seems to lack a sense of urgency to achieve this, although several NRENs are exploring this in greater detail. For virtual and research organisations there is a higher degree of motivation, perhaps driven by the fact that their researcher membership can be global, eID schemes are limited in scope and extent and not all researchers have an easily identifiable home institution affiliation.

One key reason linked to a desire for identity verification process improvement is an increase in efficiency of the vetting process and a consequent reduction in costs for the institution. The vetting process can be both time consuming and error-prone, both when conducted face-to-face and perhaps even more so with most current remote processes (e.g., video interviews with document presentation). In the face-to-face scenario staff time and costs can be high and there is also a burden on the user to attend. In the remote scenario, the presentation of documents in a live video chat session relies heavily on the capabilities of the equipment used to capture the information unambiguously. Linked to this is a fact mentioned by some that there is a lack of suitably trained staff who can verify official documents, leading to a risk of error and subsequent illegitimate approval.

Another reason cited was the growth in online learning (as a part of a specific course or a lifelong learning program), driven both by the current COVID event and an ongoing trend within the sector. This gives rise to a need to verify large numbers of individuals (mainly students) who will never visit their home institution, or will do so very infrequently.

Finally, since the enrolment process typically consists of several steps separated by considerable periods of time, in some cases there arises the possibility of using partial verification to gain fraudulent access to benefits and services for which an individual should be fully vetted. Whilst this may not be a direct problem for the institutions concerned, it could indirectly affect them through the institutions' connections with the parties providing these services and benefits.

Current status and solution landscape

Amongst the various classes of stakeholders, the study found that in many cases there was neither significant knowledge of the solutions available nor solutions in place or plans to deploy them. In general, it is still early days for identity verification solutions within R&E. For research organisations and NRENs, the real need for such solutions is unknown or sufficient pressure is lacking to investigate further. Few stakeholder organisations across the stakeholder categories have indicated more concrete plans on piloting or rolling out a service. However, awareness of the issue is rising and some have already taken the first steps, such as discussions with their member institutions or the assessment of the market including available solutions.

One of the impressions borne out by the study to a limited extent was that stakeholders who already have responsibility for identity verification (by virtue of their type or the services they run) are best placed to move ahead quickly in this area. It was also noted that even in countries where eIDAS eIDs could be used, the minimum data set provided by eIDAS would not be sufficient for identity verification within R&E.

Apart from the mentioned national eIDs and eIDAS initiatives, eduTEAMS was also identified as a possible home for identity verification. As regards third-party vendors, the following were cited (the list is similar to the one in Appendix A):

1. Signicat
2. ReadID
3. Nixu
4. Nets
5. POSTIDENT

Requirements

During the interviews, many potential requirements were mentioned by participants along with some endorsement of the requirements that were proposed. For completeness, the requirements are gathered in this section, grouped under thematic headings. Further work is needed to determine the correctness and sufficiency of the requirements and their inclusion does not imply endorsement at this stage.

Identification and verification methods

- The solution must support official photo ID documents issued by governments and be able to extract data from government-issued documents and verify its integrity. Support for other types of documents is a plus.
- It must follow identity federation rules on the use of national ID, driving licenses, paper documents, passports, and other national ID documents. Validation of documents against an authoritative source is a lower priority and would be left to the

solution providers to offer once the corresponding APIs or services are provided by national authorities.

- Any solution should support different official ID documents but also paper documents. Although the verification of these ID documents could be automated, their validation and checking should be left to the institutions; it is unclear if these documents and their audit trail should be tracked by the system.
- It should support a liveness check during remote (not F2F) interaction, i.e., that a real person is presenting the claimed identity and supporting document(s).
- For completeness, the solution could support biometric checks, primarily face recognition with reasonable spoofing protection. Biometric and automated checks would help with most foreign students without access to supported eID systems, bank or mobile credentials and ease the validation of passports to skip the help desk, but only if provided with acceptable usability. Biometrics is desirable but not suitable for high LoA (Level of Assurance).
- Should improve dealing with students who are abroad and without a verifiable bank account that could be used in identification.
- Some high-LoA services may need F2F vetting, but it should be avoided in regular education, except where necessary for specific applicant groups.
- For use by scientific communities, it must support validation of both EU-based end-users and outside the EU (e.g., US and Japan).
- Researcher's identity confirmation based on their academic record and attestation via ORCID, where they prove custody over an ORCID iD, may be needed.
- Support for additional mechanisms of identity assurance might be needed, beyond what GÉANT and current academic identity management services provide.

LoA

- Must support LoA definitions of participating identity federations (aligned with REFEDS) and keep an audit trail of provided evidence and proofing process. Should have the flexibility to apply additional checks (as needed) to achieve the necessary LoA.
- Participating organisations may ensure various levels of assurance for their users (due to a set of factors, e.g., local policies)
- High LoA is not necessary in education.
- ORCID iD can be used in support of some specific LoA validation workflows as a source for or intermediary to collateral corroboration, as ORCID provides pointers to individuals' data and academic trails which then may be confirmed with corresponding authorities, such as alma maters and employers. Moreover, in humanities, additional attributes such as publications lists may need checking to provide further identity assurance, so adjacent supporting services akin to InAcademia may be useful. This is also the case in the life sciences where checks on publication history are used to gate access to primary research data, such as genomic information.

Usability, fallback and processing speed

- A user-friendly, simple and fast process is the key. The solution should be easy to use and with a low entry barrier for onboarding users. It must be inclusive for all, taking account of disabilities and the lack of mobile device access.
- The user component should be easy to deploy by the end-user, ideally without the need to install a dedicated application. It should preferably work without a custom configuration or installation beyond a web browser, Zoom or, at most, a mobile application.
- The integration of home institutions' help desks as a fallback will be needed if the solution cannot verify identity, but its staff should see as little of personal details as possible.
- For high-LoA scientific communities, ease of use is important, but users are expected to be technically competent.
- Typical basic identity validation should not take longer than one day, but not at expense of security.
- Although speed is desirable, if identification is only done once, it is less of an issue, but depends on the use case.
- The digital process must be as user friendly and inclusive as possible.
- Institutions want an automated digital process that will reduce the workload.
- Should be as automated as possible, but needs to be able to handle exceptions.

Issuance of credentials

- Must clearly separate identity from authentication and authorisation.
- Some national federations require strong authentication when giving a password. This is done at eIDAS substantial LoA, where identity proofing and validation of a natural person typically includes F2F validation, which is normally done via help desks but could extend to banks, delivery or postal services, etc.
- Should support some means of issuing a physical or mobile-based token to allow access to premises and high-value equipment.

Integration and APIs

- Access must be provided to the verified data and the solution must integrate with the IdM system.
- A simple API (e.g., RESTful) interface to the backend and other systems is preferred.
- It would be nice to have SAML and OpenID Connect.

Trust and regulatory compliance

- Must be accurate and trustworthy.
- Privacy protection is a must.
- Must comply with the regulatory requirements for the validation of identities.

- Must comply with GDPR.
- It is particularly important to ensure that liability issues are addressed.
- In line with eIDAS, it must recognize electronic identification from all EU member states.
- Apart from being able to access other eIDAS compliant services, the envisioned solutions and infrastructures should serve as a basis for eIDAS compliant electronic signatures, certificates and trust service(s).
- Need to take account of IGTF (Interoperable Global Trust Federation) requirements for trust.
- Both the managed evidence and the system should be available for audit.

Cost and flexibility

- The cost should be low and comparable to alternatives.
- Optional features (e.g., biometrics) must not affect solution cost for small institutions. It must be flexible, so it can be offered in diverse ways. An ability to switch identity verification providers in/out of the integrated solution is important. Any service should be capable of white-labelling (obviously not needed if the solution is just a framework).
- Technical flexibility is important, as additional identity checks may be required by some institutions.

Business factors

An information exchange platform supporting actual business decisions or trials would be very welcome, particularly if it would improve decisions and increase savings. NRENs should be involved at an early stage and could set a vetting quality standard. This platform could later lead to a joint approach and its ability to enhance results and reduce prices would be apparent.

A procurement framework for identity verification is not a primary interest at the moment, as the interviewees are mostly into exploring the needs of educational organisations or getting the internal experiences by trying approaches or solutions and drafting their contractual approaches. Some have already taken this path and have already made time or scope limited commitments: one with a commercial partner, another with a GÉANT service. At some point, most would be interested in working with GÉANT and could take part in joint procurement, but it should allow for optionality and secondary local procurement for compliance with tender requirements.

Provable cost savings would bring in larger institutions and their agglomerations, while a direct involvement of smaller institutions would depend on a significant reduction of current costs or costs of alternatives. The elements of the economic justification would include improved assurance (although it is hard to quantify in monetary terms) and savings on the service desk or SMS challenges costs. Prices should depend on the institution type, the yearly number of verifications, and included options.

Due to the seasonality of student and visitor registrations, a subscription-based model would be preferred over a transaction-based one. It should be based on a fixed monthly price without monthly caps. However, some would accept a part of the cost to be per-transaction, particularly if some subjects could be charged for the service while getting a clear benefit from paying, e.g., not having to come to the central service desk.

Trust and privacy are important. In some countries, strong regulation and existing state-mandated mechanisms and structures could be a significant obstacle to novel approaches. On-premise or hosting by NRENs is preferred by most, but cloud-based solutions would be also acceptable if GDPR and other requirements were complied with. Provider's support in local dealing with requirements and security would be highly appreciated. Since available commercial solutions manage data elsewhere, the local storage of PII may not be a real option. This is acceptable if the solution provider can be audited by NRENs and their members. In addition to the more uniform legal environment, this is an argument for solutions that keep and process data within the EU. Also, some would rather avoid relying on big-tech providers, who, they suspect, could be rather interested in a market takeover, precision marketing, or consolidation of identities.

Implementation approach

Interviewees noted that implementation partnerships will depend on who is offering the solution and what data and features are being provided. GÉANT is well placed to offer a solution if it would provide economies of scale within a suitable delivery timeframe. However, the solution should allow a level of diversification, as the institutions are individually responsible for the information and the systems they use apart from the mandated baseline (i.e., commonalities such as eIDAS eIDs, where one proxy could serve different identity providers). Also, the individual institutions in charge could transfer some of their responsibilities to federations, which would be also responsible for audits against rules. All this work could be supported with a forum for the exchange of information and best practices, and easy access to available solutions and services, with information about indicative costs and integration APIs. The provided APIs must be suitable for integration.

When it comes to virtual organisations and specialised scientific collaborations, the participating IdPs are likely to continue working as they are, so these collaborations might require additional processes for individuals without an IdP and for IdPs that do not share the integrative base or have insufficient LoA.

The implementation could start with a subset of target populations (e.g., students from pilot universities, foreign major students who are starting their studies from abroad) and then roll out to all. Adoption of any centralised components would be a lengthy process, where larger institutions could be faster, while smaller ones (long-tail) would take longer as not using them would not be a major problem for them.

Some underlying services and parts of the process could be centralised, but institutions would remain responsible for the final identity verification and human labour involved (e.g., matriculation). Also, the personnel from the concerned institutions should perform any human-based remote identity verification outside e-government identification. Some service desk functions could be automated in the process, but some manual fallbacks will remain. The initial identity screening (basic identity check) and information gathering could be externalised with an agent-based or automated service offering. In the long run, a centralised framework is preferred but would have to deal with several and national systems that should converge first.

7. Conclusions and next steps

The increase in research collaborations and the rise in student mobility, coupled with the need to access digital services, present a problem for the efficient and effective onboarding of users. Automated and automation-assisted document-based identity verification could help provide a solution for the R&E community, and this study set out to investigate the interest in options for delivering such a solution or to see whether the available information is sufficient at this stage.

This work has confirmed that the problem of identity verification is most acute in the vetting of foreign students and researchers since stakeholders are content in their ability to manage identity verification for national or local users (although they would like to enhance accessibility). Although automated systems for document-based verification could bring benefits in terms of improved efficiency, lower costs and improved accuracy of identification in general, there is no strong pull from the community, at this time, for such solutions, but they are seen as an emerging topic. This is partly since the market is nascent within R&E, knowledge is limited and, in many cases, investigations of need are still to be undertaken.

The early state of the market implies that it is not possible to determine in sufficient detail what the requirements for such a solution would be and how these requirements (both functional and organisational) might map to the capabilities provided by a few commercial solutions that have been examined. Nevertheless, valuable information has been collected about these solutions and some validation of the benefits has been asserted by a few stakeholders even if, at this stage, the applicability and the full impact of the benefits to their organisation are unclear.

A small number of the stakeholders who took part in the study were either aware of some of the solutions available, or, in a limited number of cases, had engaged with one or more providers or had established pilots. This was the case across the various stakeholder types; organisations which already have responsibility for identity verification will be the leaders in investigating the applicability of such solutions to their specific circumstances.

Document-based identity verification is one of several methods that could be used to prove an individual’s identity, but it is of particular interest in the short/medium term despite the increasing availability of national and pan-European (eIDAS compliant) eIDs, efforts to support student mobility such as the European Student Card initiative and the prospect of

SSI. The study examined the current state of play of these initiatives to determine their impact on the need and utility of ID document-based systems.

Our findings indicate that there is a strong need for information about the comparative capabilities of available document-based identity solutions. The availability of such information could enable a wider discussion within the R&E community about the requirements for the community uses. Although the aims of various eID schemes and student mobility initiatives might be considered satisfactory regarding the requirement for identity verification, the rate of roll-out and adoption and issues regarding interoperability do not seem to invalidate the need for ID document-based systems, at least not in the short and medium term. Furthermore, in the case of student identity schemes, such as the European student Card Initiative, the focus is primarily on authentication, with the identity verification phase being left as the responsibility of either the home institution or government.

Further work in this area is needed for a definitive conclusion. These follow-up actions can be recommended as useful next steps:

- Establishment of a platform for information capture and exchange

  Stakeholders have reached various stages in their exploration of identity verification. A SIG of interested parties should be formed to further explore the topic and to set up a repository of information for the benefit of all. The owner of this platform should be an entity with relevance and longevity as the investigations will take some time.

- Comparative analysis of available commercial solutions

  Several commercial solutions exist for document-based identity verification and this study has listed some of these and has included some initial research. However, since many of the capabilities are broadly similar at the level of examination, it is recommended that a more detailed comparative analysis is performed which will involve an in-depth engagement with each of the vendors. This could be carried out by the group described above, or as part of an ensuing study.

- Survey of the R&E community concerning the findings so far

  The study has reached a set of conclusions based on interviews with a limited set of stakeholders due to limitations of time and resources. It is recommended that the findings be posited in the form of a survey that can be broadly distributed to NRENs and other stakeholders to confirm the general validity of the findings. This survey could be carried out by GÉANT or by the group proposed above, who would take this topic forward.

- Preliminary business case for a community-operated service

  Although the level of information available during the study was inadequate to develop a business case for a community-operated service or determine the best form for such a service, it is recommended that this be examined again after a more extensive survey of the community (as described above) is completed. Development of an initial business model canvas based on this could be helpful in evaluating the business case.

Although this study is unable to conclude on what would be the best approach based on the options initially cited, it has nevertheless collected a considerable amount of useful information that could set the future trajectory for further investigation in this area. Hopefully, if the recommendations are followed, they will be of benefit to the community and could lead to a more definitive conclusion on the matter. Based on the materials collected during this study and the prior related work (including the earlier work within the Incubator), a package with key relevant issues, requirements and checklists could be developed. It would be a result of tailoring this report and complementing it with additional information (potentially including the outputs of some of the previously listed steps, if conducted) and would serve decision-makers at various levels to refine their plans and validate ongoing activities.

References

This section contains background references used in the preparation of this document and not explicitly referenced in the text.

ITU-T X.1254/ ISO/IEC DIS 29115

NIST 800-63A (Digital identity guidelines)

REFEDS assurance framework

eIDAS regulation

eIDAS – eID assurance framework

Community-based Trust establishment

Regulation (EU) 910/2014 on electronic identification and trust services for electronic transactions in the internal market)

https://ec.europa.eu/digital-single-market/en/news/egovernment-benchmark-2020-egovernment-works-people

https://www.eurosmart.com/implementation-of-the-eidas-nodes-state-of-play/

Enhancing University Services by Extending the eIDAS European Specification with Academic Attributes, https://www.mdpi.com/2071-1050/12/3/770/htm.

https://ec.europa.eu/eurostat/statistics-explained/index.php/Learning_mobility_statistics

Regulation (EU) 2018/1724

European education area

http://roadmap2018.esfri.eu/

Appendix A – Solutions survey and assumed requirements

General requirements

The process must be easy to use in the sense that it does not generate friction for the customer potentially leading to process abandonment.

Validity checking should ensure that the integrity and authenticity of documents are checked and that:

- The document is genuine and has not been forged;
- The document has validity, i.e., it has been issued by an authoritative source;
- The document has not expired;
- The document has not been stolen, revoked or suspended.

The time required to verify a user's identity should not exceed N minutes for typical cases, as specified in advance.

The solution should strike a balance between usability (efficiency) and quality of vetting.

Verification should be as far as possible automatic; if this is not possible, escalation to a manual process should be used. This may come at an additional cost.

Documents may optionally be verified against known authoritative external sources if the document scan does not provide a sufficient level of assurance.

The process must be fully compliant with GDPR with respect to PII and image storing,

Logs may need to be made available for audit purposes.

The process must support identity documents in the EU countries as a priority, but also those from all over the world, as VOs may have users from many different countries.

The process must be capable of supporting low, medium and high identity assurance levels, as defined in REFEDS RAF 1.0.

Feature descriptions

- Technical features
    - Facial match – ID doc photo/video matching
        - OCR
        - NFC
    - Presentation attack detection (liveness detection)
    - Supported document types and geographical coverage
    - Document validity check
        - Expiration

        - Revocation
        - Check of the validity of ID docs with issuers (national authorities)
    - Service desk able to handle exceptional and high-LoA situations through human interaction
    - Available in at least 3 EU countries
    - Umbrella provider – ability to collect claims from other services (third parties)
    - Integration capabilities and available interfaces
        - Identity Access API
        - SAML authentication
        - OIDC authentication
        - Web SDK
        - Mobile SDK
        - White-label mobile application
- Commercial features
    - Setup fees
    - Subscription fee
    - Per-transaction cost
    - Educational organisation discount
    - Regulatory compliance
        - GDPR (EU)
    - Liabilities (for service malfunction, e.g., false positives)
- Operational Features
    - Max. automated verification time
    - Verification feedback – binary or annotated
    - Deployment model
        - Public cloud
        - Private Cloud
        - Hybrid Cloud
    - Support mechanism
        - Email or ticketing
        - Response times
        - Support levels
        - Reporting and transaction data
    - Data storage and security

        - Identified locations for data storage
        - Ability to control data storage locations
        - Protection mechanisms for data at rest
        - Protection for data in transit
        - ISO 27001 certified provider
    - Availability and reliability
        - Service guaranteed availability
        - Resilience mechanisms (load balancing cluster, etc.)

Solutions provider summary

Solution providers range from smaller VC-funded start-ups to large exchange-listed players. Some organisations offer a comprehensive range of identity proofing based solutions, whilst others focus on specific areas, such as biometrics. There is a range of partnerships across the organisations with some platform providers offering one or more specific technology solutions (e.g., document scanning) from partners as a part of a more integrated platform offering.

At a high-level, and focusing on the ID document-based identity proofing, most of the solutions offer broadly similar features (document capture, template matching, biometrics, etc.) although these may be bundled in diverse ways. Most solutions for identity verification consist of two main parts:

- An SDK targeted at one or more platforms (mobile (iOS and Android), web) for data collection that combines document data extraction with biometric data, such as a selfie or a video;
- A RESTful API providing access to a backend server where the verification takes place and which is invoked (generally) from the organisational server, to submit the data and retrieve the verified results.

A comparative analysis of the solutions is beyond the scope of this report but would be a key enabler for any solution selection process.

The identified solutions are listed in alphabetical order:

ElectronicID – https://www.electronicid.eu/en

ElectronicID (eID) was founded in Spain in 2013. It now has headquarters in Madrid and offices in Berlin and Paris, and an international presence in Tokyo, Hong Kong and Dubai. It has created patented video identification technology that claims to identify the user in seconds and offer the same level of security as face-to-face verification.

It offers two relevant solutions: VideoID for picture-based identification and SmileID using facial recognition for authentication. The VideoID solution records and analyses the front and back of a presented identity document and uses OCR to extract the data and validate it. A video is then captured, analysed and compared with the identity document for verification.

SmileID provides an authentication service based on facial biometrics. A web SDK is available for integration into a web client on desktop or mobile devices.

The pricing model is based on annual licenses, the cost of which is calculated depending on the solution, given that different solutions are offered depending on the security level required and features, as well as on the estimated volume of operations per year – the higher the volume, the lower the price.

iProov – https://www.iproov.com/

iProov is a London based identity authentication start-up founded in 2011, with offices in Maryland and Singapore. It is focused on 'genuine presence assurance', which ensures that an individual is an identified real person and is performing the authentication process in person, in real-time. It has secured several high-profile customers, including banks and U.K. and U.S government departments.

The organisation offers several solutions to support enrolment and biometric verification. Enroller works on Android, iOS or the web and matches images sourced via optical or NFC with an image taken of the face, and provides API for result query over REST. Face Verifier allows for authentication using a user's face against a pre-enrolled facial template. Basic Face Verifier uses liveness assurance to ensure the real person is present and the authentication event. Finally, Palm Verifier uses a contactless hand-based verification approach to check whether the user is the correct real person.

Jumio – https://www.jumio.com/

Jumio is a U.S-based company headquartered in Palo Alto, California, with global offices. It was founded in 2010 and focuses on end-to-end identity verification services through their KYX ('Know your X', where X= customer, employee, etc.), identity proofing and transaction monitoring solutions. Jumio is one of the major players in the market, having verified over 300 million identities in more than 200 countries.

Jumio has a broad range of solutions, from ID document verification using optical and NFC mechanisms, to facial recognition and video liveness checking based on ML algorithms. They also provide backend identity and address screening services and transaction monitoring through a variety of databases. They provide a complete end-to-end identity platform that ties together the solutions and allows for a seamless experience.

Keesing – https://www.keesingtechnologies.com/

Keesing is a Dutch company headquartered in Amsterdam with offices in the U.S. It was established in 1911 and originally focused on banknote identification technology. It maintains a global document database (claimed to be the world's most comprehensive), which includes details and images of thousands of identity documents, which are used as templates in the validation process. It uses a wide range of solutions for technology partners to deliver its end-to-end identity verification solutions.

Keesing offers both remote and face-to-face verification solutions. The remote solution offering supports document verification, with biometric facial recognition and liveness checking. Document checking may be accomplished (where enabled by the document) using automated optical and NFC checking or by comparison with the extensive document template database.

Mitek – https://www.miteksystems.com/

Mitek Systems is a U.S company headquartered in San Diego with offices in Europe. It was founded in 1985 and provides several solutions for mobile capture and identity verification. It is a Nasdaq-quoted company and one of the major players in this area, having provided solutions for over 6,400 organisations and a total of 80 million end users.

Mitek provides a range of solutions for identity proofing, including mobile solutions for document capture with additional facial recognition technology. MobileVerify scans an identity document and identifies it via image classification from a database stored on Mitek's servers. Data is extracted via OCR and the document is validated. Once done. biometric comparison is performed between an image of the person and the identity document. They also provide MiSnap mobile capture SDK to optimise the image capture process and MiTek NFC for authentication, also in the form of an SDK.

Mobbeel – https://www.mobbeel.com//en

Mobbeel is a small start-up headquartered in Spain, founded in 2009, which focuses on biometric authentication solutions. They have several large clients in the banking, finance, health and telecommunication sectors.

Mobbeel provides Mobbscan, a solution for identity verification using documents and MobbID, a biometric identity verification solution. MobbScan supports both a web interface and native applications for iOS and Android and supports 250 types of identity documents across 194 countries. MobScan supports document reading using both NFC and OCR and includes facial recognition and liveness checking for verification. MobbID is an authentication solution using multiple biometric mechanisms – face, voice, fingerprint, iris, including liveness checking.

Nets – https://www.nets.eu/

Nets is a Danish company focused on digital payments and related services. It is represented in 20 countries across Europe and is a large company employing 4,100 people. Its customers, primarily banks and merchants number some 250 and 700,000 respectively and it supports over 260,000 corporations with its services.

Nets operates E-Ident ID broker service that hides the Nordic government/bank MFA services behind a single SAML/OIDC proxy. In April 2020, Nets launched a new service Passport Reader, which extends E-Ident to support photographic comparison using an Android/iOS app, which scans (optically/NFC) an ICAO Doc 9303 compliant document (passport, driver's license, or residence card) and compares it to an image taken from a

user's face. The reliability of the enrolment is estimated using a False Acceptance Ratio that is provided to the client together with the passport data.

ReadID – https://readid.com/

ReadID is a product of a small organisation, InnoValor BV, based in the Netherlands and founded in 2013. InnoValor provides a range of research-based consultancy services in innovative digital technologies for the government, finance industry and service providers.

The product allows identity document verification by deploying a mobile application (or mobile SDK) to unlock and extract data from the chip embedded in the document over NFC. This can further be sent to a backend server for verification (and combined with facial matching) and the results can be accessed via a RESTful API as JSON, XML or PDF. Liveness detection via a third-party can also be added to the solution.

The pricing model consists of two components: a monthly fee for hosting, support and maintenance of the solution, together with volume-based per-transaction pricing, which depends on the functionality provided (i.e., the addition of liveness detection incurs an additional cost).

Onfido – https://onfido.com/

Onfido is a U.K headquartered company with global offices. It was founded in 2012 to allow organisations to verify their users. It has over 400 employees and a customer base of around 1500 organisations across finance, healthcare, telecommunications and others.

Onfido provides several solutions that allow identity document verification using a mobile SDK to extract data from a variety of identification documents (allegedly 4600 document types from 195 countries) together with still or live facial matching to confirm document ownership. They also provide authentication through facial matching with a previously verified identity.

Signicat – https://www.signicat.com/en

Norwegian organisation founded in 2006 and now present in around 9 European countries with over 1000 customers. Focused on the digital identity lifecycle and building the most comprehensive digital identity platform in Europe. Has grown organically and by complementary acquisition (idfy – 2019, Connectis 2020). CAGR of around 39%, with particular strength in the Nordics and a number of large corporate users within banking and finance. Processes over 3M transactions/month within Finland. ISO 27001 certified.

Signicat provides a number of identity and authentication solutions as a set of complementary services.

- An identity hub cloud platform allowing a single API to be used to connect a service provider to a broad set of European-wide eIDs (25+).
- Assure, an identity verification (scanning and verification of digital identity documents) solution for where eIDs are not available. Assure uses the services of other third parties (Onfido, ReadID and Electronic ID) for the document scanning and

liveness checking and provides a RESTful API as a single point of integration and collection of user data.

- Connect, a service for authentication and approval using a mobile device supporting both SAML and OIDC. Additionally, the Assure service can be used for establishing an identity during enrolment.
- Sign, a service providing eSignatures for digitally signing documents.

The pricing model for the Assure service is based on a monthly subscription fee and a per-transaction cost. The level of the fees is set according to the level of assurance required, the partner scanning service used and the capabilities used within that service.

SisuID – https://sisuid.com/

SisuID is a consortium led by Nixu, a Finnish cybersecurity company. They have developed an ID proofing and authentication platform. Their stated objective is to develop an open and affordable authentication platform as an enabler for business and to amortise the costs across the community and to share development.

The consortium currently has several pilots running and is trying to achieve a critical mass of customers to launch the SisuID service. However, at the time of writing, the platform release has been delayed and is not yet deployed in production due to not reaching a critical mass of customers. Nixu has promised to make the platform itself available as an open-source product so anyone (like GÉANT) can run it. The platform appears to the relying services as an OIDC or SAML Identity provider and includes an iOS/Android app for identity proofing (support for biometric passports over NFC and face video liveness check) and, once the ID proofing is done, as an MFA token for the user. SisuID makes use of external services for document and video matching.

SisuID pricing is based on an annual fixed-fee subscription model with yearly active user thresholds. The level of the subscription fee is dependent on the number of users of the platform as the costs are shared.

Appendix B – Survey questions

Initially, the team collected a number of questions:

1. Would such a service be accepted by your member organisations and how would it be propagated to them? For example, would it be decided upon and supplied centrally (e.g., by you) or it is acceptable to follow a framework agreement for NRENs/institutions to select a solution that most closely matches the collectively negotiated terms?
2. How will customers of the service select the most appropriate service? Will this be up to the customer based on previous or supplied knowledge? Will it be selected based on customers' requirements?
3. Who are the end-users or user groups for whom you manage identities?
4. How do you currently manage the enrolment and community membership of new users?
5. What are the problems in the current enrolment mechanism that you would like to address?
6. Are you any third-party solutions to help in this process that you are using or considering, and how well do they address your needs?
7. If an identity proofing solution were available to support your needs which of the following requirements should be supported:
    - SAML for authentication;
    - OIDC for authentication;
    - Identity claims based on identity documents (state which);
    - Matching of document data against the actual persons;
    - Support for several levels of assurance;
    - Ability to incorporate community definable attestations;
    - Fully automatic service without the need for human intervention;
    - Validation and handling of exceptional and high-LoA situations via a dedicated service desk (achievable scenarios and associated costs should be explored).
8. Would you be prepared to pay for such a service and what type of payment model would you prefer?
9. Key parameters for service selection (price, span, maturity, features, ease of use at end-user, technical or business level...)
10. Are there other technical or service features that you consider significant?

These questions were further refined and divided into several categories so that the time spent on each area could be better distributed. A simple analysis of the questions into open

and closed types helped to set the interview timing and extent of the question set. The questions acted as a starting point for the discussions and over time as more feedback was collected the questions were modified and regrouped to be more relevant and to reduce the time spent on areas where there did not seem to be much feedback.

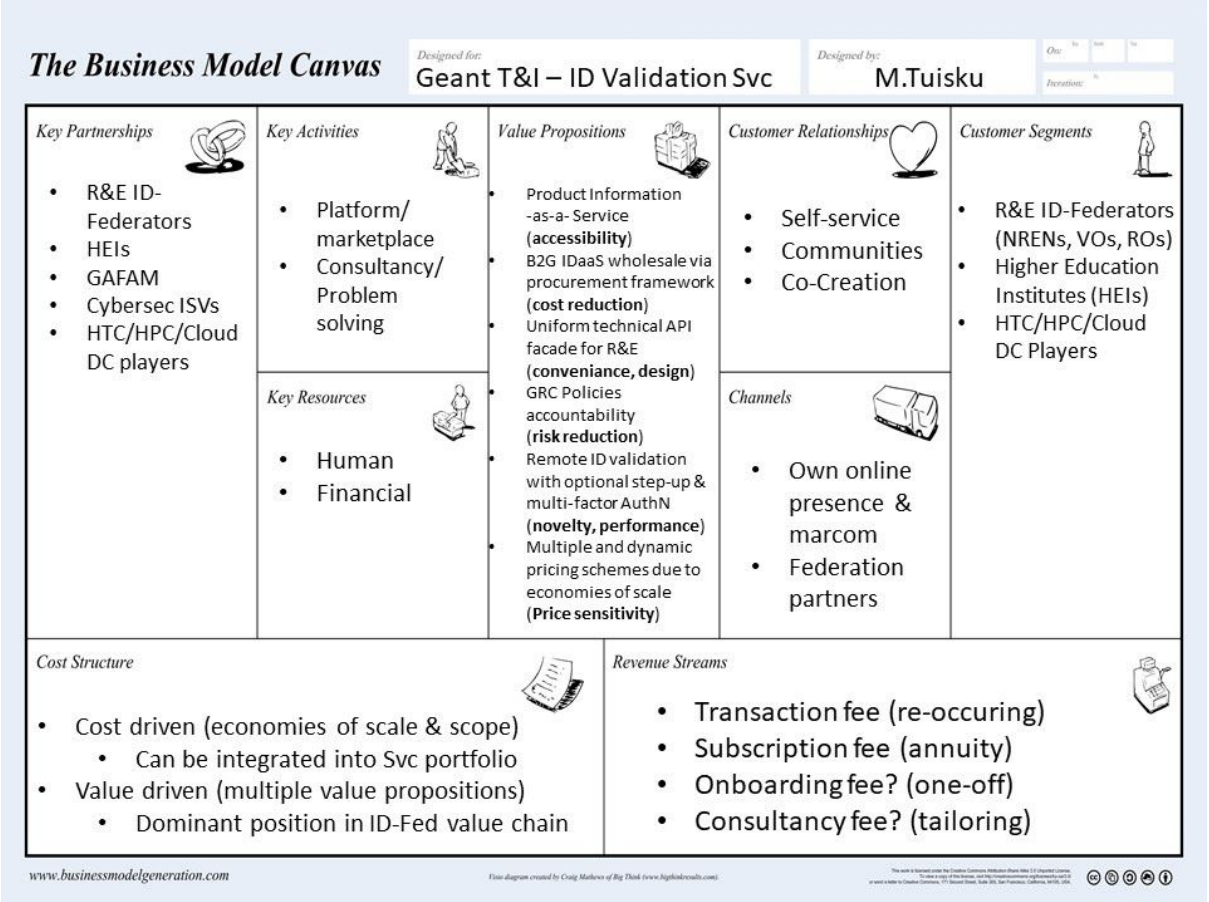
Initial interview structure and questions

- Introduction (scene-setting prelude)
    - We are trying to understand how organisations are addressing the problem of identity verification to enrol users into their organisations and if the current approaches being used are meeting the needs.
    - We are looking to see if a common set of use cases and requirements exists and how these could best be best supported for the whole R&E community.
- Present approach
    - As an NREN do you manage the enrolment of users centrally or is this devolved to the individual institutions or communities?
    - Is the reliable verification of the identities of users a problem for you or for the institutions or communities you support?
    - What solutions do you or your institutions or communities currently employ for identifying users?
    - Do you face any problems with the existing solutions being used? If so, please describe what they are.
    - What proportion of the institutions or research communities you support are affected by these problems?
- Future needs
    - To address the problems identified are there plans to develop or procure any additional solutions – if so, please describe?
    - What are the key requirements driving your approach (e.g., clear costs, the suitability of solution, QoS, availability of skilled internal resources...)?
    - What are the principal use cases that must be supported?
    - What should the geographical coverage of any proposed solution be?
    - When do you plan to make any such improvements?
- Requirements on identity strength
    - What forms of verification do the various relying parties need to support?
    - What types of identity and supporting documents should be supported?
    - Do you have a need for identity documents being validated against authoritative issuing sources or are authenticity and integrity checks sufficient?

- Is person to person (remote or F2F) vetting a necessary requirement for some of your assurance levels?
- Is automated matching the ID to the actual person (liveness and biometrics) acceptable or is an in-person check always needed?

- Technical requirements
  - What interfaces to your existing AAI/IdM backend system are required?
  - How important is it to your users to have a solution that would work with minimal technical knowledge (e.g., simple web forms, and in-person or telephone contact vs installing an application, configuring the OS, using command line…)?
  - Do you need to combine several identity assertions?
  - Should the solution support authentication of users as well as identity proofing?
  - What authentication protocols should be supported between the solution and the relying party?
  - What is the maximum time for verification that would be acceptable to your users?
  - Do some relevant scenarios require delegation of identity confirmation to community members (Web of Trust, P2P...)?
- Business needs
  - Is there specific help you need in identifying the best solution for your needs?
  - Would you prefer to manage the procurement of any solution yourself, or would you prefer this to be done centrally (e.g., by GÉANT)?
  - Would you prefer any solution to be centrally operated within the R&E community (e.g., GÉANT)?
  - Would a solution or PII data have to be on-premise due to regulatory requirements (e.g., ISO 27001, GDPR) or local policies?
  - Would you be prepared to pay for such a solution and what cost would be acceptable for this type of service?
  - What contractual approach would work well for your organisation?
  - What types of pricing model would be acceptable to you (e.g., flat-rate yearly subscription, per-transaction fee etc.)?
  - How do you think the risk should be shared between the other solution provider and your organisation?

Appendix C – Business model canvas



Initial Business Model Canvas