21-11-2019

Distributed Vetting - Multi-Factor Authentication Report

Authors: Branko Marovic (AMRES/UoB), Jule Ziegler (DFN/LRZ), Mihály Héder (SZTAKI)

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).

Table of content

1 Introduction

The GN4-2 eduTEAMS project conducted a pilot with the deployment of a Step-Up Authentication Solution for the LifeScience community. Many of the research communities eduTEAMS supports are highly distributed over Europe and beyond. One of the challenges this pilot therefore faced was how to deal with the vetting of identities and binding of additional authentication factors in a distributed way. This activity investigates how token registration in research can be scaled for scenarios where participants are distributed over the EU and beyond.

In this report we investigate and describe existing methods and review the related work, focusing on R&E but also touching on other sectors. We identify those methods that can be readily implemented or adapted for use in a distributed environment. Since the Level of Assurance (LoA) of Two Factor Authentication (2FA) solutions rely heavily on the quality of the performed identity vetting, token registration and binding, different options and dimensions need to be considered. So various methods, token types and LoAs may be needed by different groups. Ideally, the outlined approach should be able to accommodate varied flows and data associated with diverse communities, needs and authentication technologies.

The available options are different in many dimensions. For instance, one important dimension is the type of the authenticator to be used, which in turn is established by the value of the defended resources, potential impact of a successful attack and the likely risk of high-sophistication attackers. High risk or impact may justify the usage of dedicated hardware authenticators. However, in this case, the tokens need to be distributed to geographically dispersed users, on top of which identity vetting may become just an additional and quite a marginal cost (e.g. by using a delivery service that mandates checking the ID of the recipient, as many mail services do).

As a counterexample, in less demanding cases soft tokens, like smartphone-based solutions might be sufficient. With no need to physically distribute the tokens, if we require physical contact for the vetting, all the cost is generated by the vetting method.

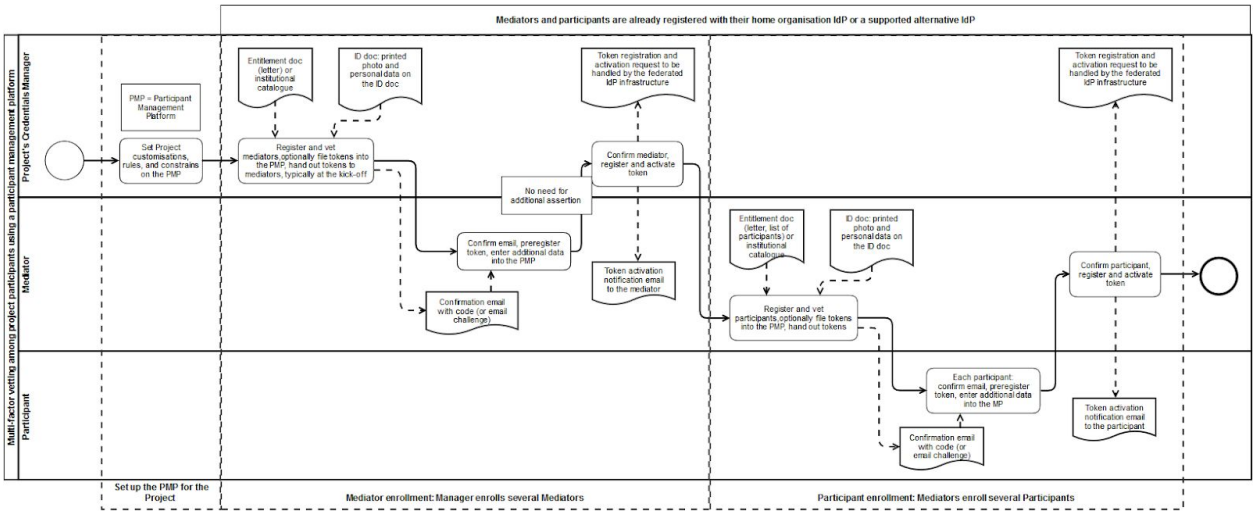
Another dimension is the cost as a function of organization scale, geographical distribution and growth. For instance, in the case of a single physical registration desk, the marginal cost is somewhat lowered with each additional new vetting done, since the capital expenses, like the cost of training and office outfitting, are distributed across many vetting actions. And yet, the function may not be sloping downwards as expected, since typically an organization or research community will grow on its edges and the newest members may come from further away and their travel cost may increase the costs more than they are decreased by the volume effect. With other methods, like peer vetting, the vetting network may grow together

with the community that helps keep the travel costs low, but this may entail some tradeoffs on the security side. A larger community may entail more stringent rules and higher assurance thresholds, at least for more sensitive roles.

This also shows that, besides the type of the authenticator and the eventual scale of the organization, the growth dynamic is also a factor - it may be affordable to do the vetting once-and-for-all for a project by a task force but not feasible to maintain a permanent capacity to do that.

Just like with the idea of compounding the physical delivery with ID checking by the postal delivery person, we may compound the face-to-face project meetings that are already happening regularly anyways with regular vetting sessions. Therefore, this is another dimension affecting the cost: the pattern of meetings for a project, like all-hands, smaller regional meetings, etc.

Finally, there are typically several tiers of trust in a project; similarly, there may be several levels of vetting in it. In other words, cheap and expensive distributed vetting methods may be mixed or may be used in isolation depending on the given project. The illustrative flow provided below shows an example of how identity vetting, token delivery and binding may be combined for highly distributed research collaborations relying on a community-based approach. The credentials manager of the project configures the participant management system, most likely a web application, and then enrols several mediators. After obtaining tokens (or other authenticators) for them and other participants, e.g. during the project kick-off, mediators enrol other project participants, to whom they pass their tokens. All participants are supposed to be already registered with their home organisation or other supported IdP. While the mediators are enrolled during an early face-to-face meeting, other participants, for example, those who could not attend the meeting or were not yet part of the project/collaboration, may be enrolled at a later occasion. Before registering the participants and issuing the tokens, the mediator may check their picture ID documents, examine the digital trail, have a video call, make phone calls using their publicly listed work phone numbers or get a confirmation from participants' organisations or other participants from those organisations. Doing some of these may save the mediator from having to visit the participant at their workplace; however, these actions should be recorded within the participant management system, in line with the rules set up by the credentials manager. Some communities may even decide to allow remote enrollment of mediators.



Vetting of project participants and registration of authentication tokens using a community-based approach [GN4-3 WP5 Projects Vetting Flow]

The scope of this report includes research and education collaborations at the EU level but possibly on a global scale. In this case, the following assumptions apply: there is at least one early meeting, such as a kick-off at the very beginning, with several senior participants coming together from different countries. Project participants are rarely all simultaneously present; new personnel is added over time but at a modest pace. New participants are rarely added to the project team without meeting at least once face-to-face with a more senior member; the collaboration is based around a subdiscipline or shared interest, at least within the project subgroups; and there is often a tier of students or outside researchers who are not part of the project itself but use the resources and therefore may need to be granted access.

The assumptions above may help in finding low-cost but suitable solutions for distributed vetting in an academic environment - of course, all the way to reflect on the fact that some of these assumptions may not be true for a given project.

2 About this document

This document was created as part of the GN4-3 WP5 Trust & Identity Incubator (Task 2). The T&I Incubator aims to develop, foster and mature new ideas in the Trust and Identity space in Research and Education. The Incubator will investigate new technologies that currently have no place (yet) in the services ecosystem of the GÉANT project. This may include to test and experiment with potential new features for existing GÉANT services. In addition, business case development for potential new services, and developments that would improve data protection and privacy aspects in services or software are also in scope.

This document is not intended to serve as an official milestone but could be used as a public informative report about the conclusions of the **Distributed Vetting - Multi-Factor Authentication** activity of the first T&I Incubator cycle.

As motivated in the introduction (section 1), this activity investigated different approaches on how to vet identities and register tokens for second-factor authentication in research communities that are distributed over the EU and beyond. As part of this, the document describes in the following section 3 the derivation of generic actions to describe the procedure of identity vetting and factor binding in a technology-independent way. Section 4 examines and discusses existing vetting methods also taking into account the cost model. Section 5 provides recommendations for further steps and finally gives a conclusion of the activity results.

3 Taxonomy of actions

This part of the report describes the commonalities in remote vetting procedures. We call these **generalized functional units** of which remote vetting processes can be assembled. These are almost atomic actions that are abstracted away from the remote vetting procedures described in "*Remote vetting for SURFconext Strong Authentication*" (short: the SURF report) [SURF Report] and from some other remote vetting procedures that we have investigated.

For a consistent and precise-enough language, we have adopted the concepts of **ITU-T X.1254/ ISO/IEC 29115** recommendation "Entity authentication assurance framework" [ITU]. This standard is both publicly available, influential and deals with identity issues.

The following generalized functional units serve to design and implement the identity vetting scenarios for second factor and multi-factor authentication that fulfil some of the ITU-T X.1254 entity authentication assurance framework processes. Here, we are referring to the following processes from "**8.1 Enrollment phase**":

- 8.1.1 Application and initiation
- 8.1.2 Identity proofing and identity information verification
- 8.1.3 Record-keeping/recording

"8.1.4 Registration" is omitted as it is related to (later) use of the services or resources.

Of all processes described in "**8.2 Credential management phase**" - only some are addressed here, as they are related with initialisation and issuance of the authentication factors, which, in our scenarios, are closely tied to identity proofing and verification:

- 8.2.1 Credential creation
    - 8.2.1.1 Credential pre-processing
    - 8.2.1.2 Credential initialization
    - 8.2.1.3 Credential binding

- 8.2.2 Credential issuance
- 8.2.3 Credential activation
- 8.2.7 Record-keeping

The names and descriptions used in our elaborations aim to be mappable to those processes and be terminologically compatible with ITU-T X.1254 and its definitions of terms. An additional particularity for the listed processes is that ITU focuses on credentials (sets of data supporting identity or entitlement claims), while our scenarios are focused on authentication factors (something specific that is usually possessed, known or inherent). The subject entities are referred to as applicants, who are the physical persons whose identity is to be authenticated.

The below descriptions use a slightly shifted terminology; for example, they primarily refer to factors, not credentials.

Actions are grouped into four sections: **Common Actions** and three general phases (**Initiation, Verification, Binding**) that are usually performed in a sequence.

Descriptions of actions are process and flow-oriented, not data-oriented. Inputs and outputs descriptions are therefore rather high-level and informal.

3.1 C: Common actions

The actions listed here are common actions that may be used multiple times at different stages and for various purposes.

C_USE_EXISTING_FACTOR: Authenticate Using Existing Factor

> The applicant authenticates with his/her factor(s) already in place and functioning in the system. Username/password login is typically the first existing factor that is readily available.
>
> This action may be used for multiple purposes:
>
> Perform authentication with the existing factor(s) to prove knowledge/possession of the respective factor(s).
>
> This action may also be used for checking the applicant's eligibility (see C_CHECK_ELIGIBILITY) based on the credentials used (e.g. by searching for a name or email address in an LDAP directory) or the attributes (e.g. affiliation) which are sent in the authentication response.
>
> **Input:** Credentials (e.g. username/password combination, certificate)
>
> **Output:** Authentication successful (yes/no), attributes if needed (e.g. affiliation)

C_SELECT_NEW_FACTOR: Select Type of Factor to be Introduced

The applicant selects the type of the new factor to be introduced if there are several options. The offered options may depend on the place of use, for example, a wider set of options may be available during initiation than with a particular subsequent (and possibly more specific) vetting and verification, where these choices may be limited.

There may be different factor types, e.g. something you know/have/are, the applicant can choose from as well as multiple realization options/products per factor (e.g. YubiKey, Google Authenticator).

Input: List of possible factors

Output: Factor selected/assigned and known by the applicant or in possession of the applicant

C_USE_NEW_FACTOR: Use Introduced Factor

Usage of the introduced factor may serve multiple purposes at different stages - to test its functioning, prove knowledge/possession/inheritance/... or to make sure that factors used in different steps of a flow match.

Input: New factor

Output: Usage of factor took place (for various purposes)

C_CHECK_ELIGIBILITY: Check Eligibility of Applicant

Check if the applicant is eligible to request an additional factor. For example, if there are some policy or contractual restrictions. Is the applicant associated with a participating organisation and eligible for a potential delivery (if offered by the vetting service) of an additional physical factor such as a token?

It may be done by manual or automated check of a directory, federated identity, or examination of a written institutional certificate.

Input: Applicant's identifying information

Output: Decision: eligible (yes/no)

3.2 I: Application and initiation

This phase deals with the initial request for an additional authentication factor during which vetting arrangements are made.

The purpose of this phase is to control access to the costly or limited resources that are used during Verification. First checks can be performed online and the user may preregister the factor which is to be introduced. Also, some actions performed here may be repeated in Verification with a higher level of assurance. Forcing the applicant to perform them in Initiation may reduce the later friction and disputes and trains the applicant for faster interaction with the personnel involved in Verification.

Skipping of Initiation may be allowed by the service policy to facilitate access for the users who need direct support. It may also be abolished for all if queueing or labour costs during Verification are not the primary issue. If so, initiation actions can be integrated with face-to-face vetting sessions, which makes the actions needed to link this phase and Verification unnecessary.

The actions that are specific for this phase are I_SUPPLY_FACTOR and I_ARRANGE_VERIFICATION. In some cases, this phase may be initiated by the Applicant's organisation, not directly by the applicant.

Some actions from *Commons* are referenced here, with supplementary explanations of their practical use in the “Application and initiation” phase. For the reasons explained above, all actions performed in this phase are optional, as omitting the initiation phase may simplify the enrolment; initiation-related actions are then performed during Verification.

C_USE_EXISTING_FACTOR (optional) ***DEFINED IN C***

> Optional. Used to provide both information about user identity and initial proof (with presumably weaker assurance) that the applicant is who (s)he claims to be.

C_CHECK_ELIGIBILITY (optional, requiring C_USE_EXISTING_FACTOR) ***DEFINED IN C***

> Optional. Used to check whether the applicant is entitled to request the additional factor at the moment of application, as this may incur some costs or use of resources.

C_SELECT_NEW_FACTOR (optional) ***DEFINED IN C***

> Optional, if there are several options for factors that may be offered at the start; it may affect the options to be used during the vetting phase.

I_SUPPLY_FACTOR (optional) **Supply New Factor**

> Optional, the applicant may either use his/her own factor (typically a token), get a factor assigned or buy it from an external provider. In the latter case, the applicant may also provide the payment information and delivery address. This may even be as simple as a mere redirection to an external supplier.
>
> If needed, this includes physical sending of the factor and delivery period warranty so that I_ARRANGE_VERIFICATION could be performed (unless C_USE_NEW_FACTOR

is required before it). Otherwise, it may involve initialisation of a mobile authenticator application or whatever is required to do so that it could be used during the rest of the scenario.

Input: Already collected data useful in arranging the supply, as the applicant name or selected factor

Output: Latest expected delivery date, optional deliverer and, ultimately, delivered factor, actual delivery date, and delivery confirmation

C_USE_NEW_FACTOR (optional) ***DEFINED IN C***

The factor to be registered is used, which confirms usability and possession by the applicant. This may be mandatory if the applicant is expected to possess the factor at the time of application and initiation; alternatively, this can be done later.

I_ARRANGE_VERIFICATION (optional) **Arrange Verification**

This is the optional detailing of vetting between the applicant and the verifier. Email, initiation application or other channel is used to communicate a code, appointment details or other relevant information. May include several steps:

- Creation of a (secret) code to be used at the start of the vetting procedure to identify the vetting request or the new factor used during initiation (C_USE_NEW_FACTOR**)**.
- If email is used for vetting arrangements, get the applicant's email address (e.g. from the IdP account data) or from the applicant.
- Optional location selection and/or scheduling of the vetting appointment, only if the load or the policy of the service (desk) require this.
- Provide vetting details over email or through the application, with code in text or QR, email validation link, instructions, vetting application link, service desk contacts, address and appointment details, and whatever else is needed.
- Optional email validation, if an email is required for further interaction, and if a valid email address is not already accessible and assured/guaranteed from the IdP data provided upon the previously performed login with the existing factor (C_USE_EXISTING_FACTOR).

Input: Information about the applicant factor type and factor instance (if it was available and used) or planned delivery, and applicant preferences/choice for the proofing and verification phase

Output: Appointment, code, confirmation data and instructions for the applicant, database record on the appointment

3.3 V: Identity proofing and information verification

Do the actual vetting by proofing the applicant's identity and verifying (or vetting) identity information. This may be performed by an external organisation or a separate internal service.

V_COMMENCE (optional) **Begin Vetting (possibly by accessing and validating the prior request)**

Set up the context for *identity proofing and information verification* by linking prior *application and initiation* or performing it if has not been done. Verify, resume, and potentially update the context established during the initiation, or do the key work that that is in it. For example, if the applicant is allowed to come to a service desk without prior registration, C_CHECK_ELIGIBILITY that is often done during initiation still must be performed; this may also be necessary if some time has passed since the initiation. Other initiation elements related to scheduling of the appointment or linking of initiation and vetting, such as (secret) code creation are pointless, as the applicant is already present and available for vetting.

- Vetting may be rejected and the applicant turned back if the applicant is not eligible (anymore) or if the queue is too long or the necessary resources, staff or involved key services are not available at this point.
- Restoring of the information and context established during initiation may include C_USE_EXISTING_FACTOR or use of previously created code to identify the vetting request or the factor used during initiation.
- If the validity of the email address is considered significant, a code or link may be used to make sure that the applicant's email is valid and can be accessed by him/her.
- Setting up of the context of the applicant's request may be done by restoring it after the applicant, service or desk operator uses the code issued during the initiation. The code that links the applicant with the original application is particularly useful when the applicant does not possess or know the first factor (which may require V_CREATE_DIGITAL_IDENTITY) and is not able to perform C_USE_EXISTING_FACTOR.
- If some time has passed since initiation, it may be necessary to perform C_CHECK_ELIGIBILITY again, as the applicant situation with her organisation may have changed in the meantime. This check could be done based on performed C_USE_EXISTING_FACTOR or verbally provided identifying information, which, in the case of human-to-human interaction may be a softer start of vetting than to immediately demand V_PRESENT_PROOF.

Input: Information about the appointment (e.g.link or code)

Output: Restored or established information about the applicant, appointment and factor, or rejection of further actions.

V_CREATE_DIGITAL_IDENTITY (optional) **Create Digital Identity (for applicants without first factor)**

Only if the applicant does not already possess the IdP identity (weak i.e. identity that can only be attested with the first factor). This is optional and often prohibited or discouraged and avoided except for those in need of assistance or VIP individuals. At a service desk, this is to be done before V_CHECK_PROOF to allow parallelism; it should be undo-able if any of the checks before V_RECORD_CHECKS fail. This includes checking of the alignment with the enforced policies, informing the applicant about the rules associated with the use of the created identity and associated factors, creation of the username and the password, and providing the applicant with them.

This action should be invalidated if any of the following enrolment actions fail during both vetting and binding phases.

Input: Applicant data needed for the IdP

Output: Digital identity created at the IdP, possibly with a “limited to verification” or “remove if not completed after ...” flag, so that the applicant can use it during the rest of the process, while also ensuring that the identity will be limited or deleted if the checks fail

V_PRESENT_PROOF Presents Proof of Identity

The applicant presents proof of identity, typically a sanctioned type of picture ID doc with demographic and biometric data.

C_SELECT_NEW_FACTOR (optional) ***DEFINED IN C***

Optional - change of the factor the applicant has already applied for is quite unlikely but this may offer some flexibility by modifying the original choice made during the initiation.

V_HAND_OVER_FACTOR (optional) **Hand Over Factor (to the applicant)**

Done if the physical factor is immediately provided during this phase, e.g. by the service desk. Like I_SUPPLY_FACTOR, it can also include an immediate monetary transaction. Recording of handover is probably unnecessary, as the service/operator is in the possession of a proof (obtained with V_PRESENT_PROOF) until the applicant completes C_USE_NEW_FACTOR.

Input: Available factor, potentially money covering the costs

Output: Factor handover record

V_CHECK_PROOF (Local) Check of the Presented Proof

This is a detailed (local) check of ID validity and its match with the person of the applicant. Compare the claimed identity (information) which is transmitted by the applicant or system with the applicant's identity proof and the actual person.

Read and inspect the ID doc, compare the name with the vetting request, check ID security features, optionally electronically read the ID doc, compare photo/biometrics match with the person.

For online mechanisms, a separate liveness check may be needed to match with the real-world person.

V_EXTERNAL_CHECK (optional) **Check External Validity**

Check user's identity proof (e.g. national ID document, employee ID card) against its source (such as issuing authority) or an official register for validity.

Make sure the identity proof is not expired, revoked or invalid.

Input: User's identity proof

Output: Verified identity proof

Effect on LoA: Higher LoA should require this action

V_CHECK_LIVENESS (optional) **Perform Liveness Check**

In case online identity vetting mechanisms are used (such as video identification, online document upload) a liveness check may be performed to prevent fraud. Otherwise implied by V_CHECK_PROOF it is already conducted with the user.

Example 1: Show ID document besides the head to prove ID document and holder match.

Example 2: Upload ID document and real-time recorded selfie/video which may be required to contain a response to a challenge.

Input: Any means that show liveness

Output: Liveness verified

V_RECORD_CHECKS (optional) **Record Identity Proof**

Optional record for audit purposes is produced, for example, by recording the last 6 digits of the number of the presented national ID document, but may also include other evidence on performed checks. However, the verifier should avoid the recording of excess personal data, including photos of the person or ID document.

A record should be made even after the failure of some of the checks. If any of the checks verification failed, V_CREATE_DIGITAL_IDENTITY may need to be undone, and the applicant may be requested to return the physical factor provided in V_HAND_OVER_FACTOR or even I_SUPPLY_FACTOR.

Input: Identity proof

Output: Identity proof record

C_USE_NEW_FACTOR (optional) ***DEFINED IN C***

This step is for making sure that the applicant knows/possesses/inherits the new factor and can use it. Or in case of preregistration making sure that all performed actions involving the new factor were with the same instance of the factor, as it will be bound to the applicant's digital identity. This step should be performed by the applicant and therefore may require some time to complete, and thus it could be done by the applicant in parallel with V_CHECK_PROOF, V_EXTERNAL_CHECK and V_CHECK_LIVENESS. It may be preceded with C_USE_EXISTING_FACTOR if it has not already been performed. This may be omitted if C_USE_NEW_FACTOR was done during initiation.

This may be required if it includes personalisation of the factor.

V_CONCLUDE Conclude Verification

If the entire verification was successful, the concluding record for the following *factor binding and activation* phase is produced. This factor will be bound and activated after all prior steps of identity proofing and information verification are completed with success, so the record about it and its association with the applicant's digital identity is saved at this point. Since different entities/providers may be responsible for verification and binding, and the verifier is typically engaged by the entity responsible for the long-term identity and credential management (IdP), the verifier may do it through a notification or by submitting the corresponding entry/record into the IdP system. Also, the verifier can make an internal record about the creation of this entry, possibly with details about performed C_USE_NEW_FACTOR.

Input: Success and details from other steps of the verification phase

Output: Data needed for binding - digital ID of applicant, confirmation of verified identity, factor information

3.4 B: Factor binding and activation

This phase is dedicated to the establishment of an operational link between the identity of the user and factor.

B_FACTOR_AND_ID Bind Factor to Digital ID

Create a long-term binding between the introduced factor and the digital ID of the user based on verified user identity.

Input: Digital ID of applicant, confirmation of verified identity, factor information

Output: Binding between digital ID and factor

B_ACTIVATE (optional) **Activate Binding of Digital ID and New Factor**

Activate the binding of the digital ID of the user and the new factor. The new factor may need to be unlocked, enabled or otherwise modified so that it can be used in regular authentications. For example, it may be in a state in which it was personalised and populated with all needed data but still marked as "not activated", which allows authentications with target services to fail even without contacting the factor issuer or IdP.

Input: Binding between digital ID and factor

Output: Decision: activation successful/unsuccessful

B_CONFIRMATION Inform User about Factor Activation

Inform the user about the correct/incorrect activation of the factor.

In case the factor activation was successful, the applicant can now authenticate using more than one factor.

Input: Result of factor activation (positive/negative)

Output: Message to the applicant

3.5 Detailed attributes

There are other candidate attributes that can be elaborated for the above-listed actions in general or with their specific implementations:

- Likely to be mandatory in MFA: (yes/no)
- Risks if omitted: (mostly security-related)
- Effect on the level of assurance: (how it increases or decreases the LoA)
- Other technical concerns/issues
- Potential organisational and legal (intellectual property, NREN, GEANT) concerns/issues
- Potential end-user related concerns/issues (e.g. usability or accessibility)

- Additional implementation of specific options or constrains

4 Vetting methods

4.1 Commonly available methods

Remote vetting for establishing a second factor or similar credentials is already an industry practice in many business domains. These range from live video sessions to the reading of chips embedded in passports to selfies with a driver's licence.

This non-exhaustive list enumerates some existing vetting procedures. These include predominantly national procedures identified during the incubator activity that were assessed for suitability with cross-national scenarios.

- **eIDAS compliant services**
  eIDAS [eIDAS Regulation, eIDAS Implementing Acts] is an EU Regulation in effect since July 2014. It has been developed and put in place to facilitate the Digital Single Market agenda of the European Commission.
    - https://ec.europa.eu/digital-single-market/en/trust-services-and-eid
    - https://en.wikipedia.org/wiki/EIDAS

  Its implementation acts set out detailed concepts for assurance and minimum specifications to how compliant systems should work.

- **Student cards**
    - EU: https://europeanstudentcard.eu
      The European Student card allows the registered terminals it is compatible with to communicate with their home educational institution and retrieve data. Such a card may be used in a remote vetting process.
    - Serbia (in Belgrade): http://sc.rs/sc/studentska-kartica/

- **ICAO eMRTDs biometric passports (ISO/IEC 14443 application MRTDs profile)**
  Biometric passports or e-passports are common in many countries in the world. The implementation of such passports is binding to any Schengen member states in the EU. Since e-passports may be readable by an app, they may be used in remote vetting.

  Described in https://www.icao.int/publications/pages/publication.aspx?docnum=9303, particularly in:
    - Part 9: Deployment of Biometric Identification and Electronic Storage of Data in eMRTDs
    - Part 10: Logical Data Structure (LDS) for Storage of Biometrics and Other Data in the Contactless Integrated Circuit (IC)
    - Part 11: Security Mechanisms for MRTDs
    - Part 12: Public Key Infrastructure for MRTDs

- **EUGridPMA**
  EUGridPMA is a trust network for research e-Infrastructures that is based on certificates. They have developed a policy for face-to-face vetting performed by peers, see:
  - https://www.eugridpma.org/guidelines/1scp/1SCP-vetting-f2f-0.2.pdf
  - https://www.eugridpma.org/
  - https://www.eugridpma.org/members/worldmap/

- **National eID documents**
  - Serbia, biometric ID card, with optional (nationally) qualified PKI certificates: http://www.mup.gov.rs/wps/portal/sr/gradjani/dokumenta/licna%20karta

- **Bank accounts and cards**
  Usage of own or organizational bank account (usually small transaction including some reference to vet identity). Use reference for linking.

- **Health insurance cards**
  Most people would feel uncomfortable with them in applications unrelated to health care, using health card for identification purposes may be prohibited by local legislation.
  - Serbia, with demographic data and cryptographic functions on the card: http://rfzo.rs/index.php/osiguranalica/ekartica

- **Commercial electronic certificates for natural persons**
  They exist in many countries, but they usually have very low penetration. An example from Serbia (https://www.euprava.gov.rs/pomoc/elektronski_sertifikati?alphabet=lat):
  - Post of Serbia http://www.ca.posta.rs/postupak_izdavanja.htm
  - Chamber of Commerce and Industry of Serbia http://www.pks.rs/Usluge.aspx?IDUsluge=4&t=2
  - Halcom http://www.halcom.rs/rs/proizvodi/sertifikati/sertifikati-2/

- **Postal service**
  - Post of Germany, 'postident' https://www.deutschepost.de/de/p/postident.html, with different vetting methods provided, e.g. face-to-face in a post office, video chat, German eID check
  - DPD ID check
  - UPS ID check

Remote vetting with some of the above-listed procedures and credentials may be supported with a number of general or custom video chat applications. If a standard application is used, the missing ID check functionality may be supplemented with a separate mobile application. There are many such mobile applications that are primarily demos intended to attract potential customers to custom development. Here are some examples from common app stores:

- ReadID - NFC Passport Reader
  https://play.google.com/store/apps/details?id=nl.innovalor.nfciddocshowcase
  https://apps.apple.com/us/app/readid-nfc/id1463949991
- NFC TagInfo
  https://play.google.com/store/apps/details?id=at.mroland.android.apps.nfctaginfo
  https://play.google.com/store/apps/details?id=at.mroland.android.apps.imagedecoder
- Regula Document Reader
  https://play.google.com/store/apps/details?id=com.regula.documentreader
- NFC Smart Card Info
  https://play.google.com/store/apps/details?id=com.inoapp.cardinfo

4.2 Additional mechanisms

These systems usually do not provide a formal vetting procedure but still provide a high level of confidence, based on the consequences of false identification or the web of trust within the community.

- ORCID, https://orcid.org/, does not provide any actual identity vetting or verification, and registration is trivial, but it is very useful for connecting publishers with their authors within the services they provide.
- PGP and OpenPGP:
  - https://www.openpgp.org/
  - https://tools.ietf.org/html/rfc4880

4.3 Vetting organisation and costs

In this part, we are going to discuss the methods for Identity Proofing and Identity Information Verification, considering the options by price range. Of course, there are other important requirements such as technical complexity or availability of human resources which need to be addressed as well; but these are explicitly not part of this report and could be investigated in subsequent incubator activities.

The technique for identity proofing partially depends on the security factor we wish to use. Namely, any physical factor sent out by the party that is doing the vetting, like a physical encryption device, needs a chain of physical contacts for reliable and attestable delivery. Therefore the cost of the vetting itself might be only a small addition to the price of delivery.

Hence, it makes sense to distinguish between two major categories. If there is a need - derived from security requirements and the value of the asset to be accessed - to pass on a physical token, then, presumably the delivery method generates the major part of the cost. In practice, a token could be passed on to the user by a face-to-face meeting or via postal delivery. Both methods can be complemented by checking the identification cards and papers of the user.

The other major category is when the second factor to be used does not need to originate from the organisation that is doing the vetting. Note that this can still involve the usage of dedicated hardware authenticators that may be purchased and owned by users or their organisations. Here, the registration of that key is the main issue. However, more probable is the case when the second factor is a smartphone, an additional identity provider, a mobile number for SMS tokens, and other digitally distributable means of authentication. This second-factor category does not demand a face-to-face or postal contact with the user, and therefore if we still demand it for identity verification, then all the cost is generated by the verification step itself.

4.3.1 Face-to-face vetting methods

Central registration desk

The baseline method for the face-to-face vetting and factor distribution is a central registration, where users visit the central office of an organisation to verify their identity and potentially get the needed physical authenticators. The marginal cost of vetting one user with this method is probably the highest among all the possibilities for any distributed collaboration. The cost of travel to a fixed geographical location from an unspecified origin is typically several hundreds of EUR. At the same time, this requires a high level of awareness and training of the verifier as they may need to be familiar with different types of identity cards and papers issued in various countries. Checking the validity of such documents might be challenging without access to national registries.

Distributed face-to-face vetting

The central vetting approach as described above might also be distributed to several geographical locations. Especially if there is a high number of users, the usage of local registration points may be economical. Alternatively, we may rely on peer vetting methods, where already vetted users do the vetting of nearby peers. In the case of large international collaborations, there is usually at least one face-to-face meeting in a year for the more senior participants at a minimum. The vetting of these participants may happen face-to-face, and then they can manage the vetting of others nearby to their geographical location. An advantage of this method is that there is a better chance of the seniors to be familiar with the identification used by the nearby colleagues.

4.3.2 Additional third-party identity vetting methods

Piggybacking existing delivery and money transfer services

Realistically, face-to-face vetting in an academic community enables more to vet the knowledge and background of the person than its identity. The reason for this is that the participants of an academic collaboration rarely have any access to a database where identification cards or papers might be verified, nor do they have the knowledge to recognize fake paperwork. This could be alleviated by providing the authorised vetters with an

application that would verify the identity documents in the corresponding registries. In the meantime, if this kind of assurance is needed, then it is probably better to rely on a trained, equipped and authorised postal delivery person or counter clerk to check the identification documents.

Usage of postal delivery firms, like UPS, DHL, or others, to pass on hardware authenticator or secret credential with recipient identification is one option with the cost tag of about 50€. Usage of a firm like Western Union to check the identity cards of a recipient may be significantly cheaper while allowing the passage of digital information only.

Access to bank accounts

The EU 2019 Payment Services Directive 2 (PSD2) defines organisational and technical measures that must be implemented by banks. It includes two-factor authentication and requires banks to enable trusted third parties to access their customer data, upon the customer's consent. In PSD2 terminology, these service providers are called account information services (AIS).

A service registered as a trusted service provider could leverage identity verification using the information from the strongly vetted user's bank account. Such a service does not need the financial data of the user but instead just those required for identity vetting in MFA authentication scenarios. This could be enhanced even further by using the bank's login interface for the regular MFA or at least confirm the applicant's identity during vetting. Some countries already make such a login possible [idin]. The PSD2 extends this to the EU level.

The respective service providers must have a national licence or one issued by another EU supervisory authority. The PSD2 regulates the supervision of third parties, encourages the use of APIs to provide the data and enforces a common communication protocol.

It defines that banks must provide Trusted Party Providers (TPP) access to accounts. Although TPP definition does not mandate an API, there seems to be an ongoing convergence. The NextGenPSD2 API specification is a common framework defined by the Berlin Group [Berlin Group] that has already been adopted by more than 3000 banks across Europe.

This specification includes the following functions:

- 7.4.1 Establish consent transaction
- 7.4.2 Get account information transaction

These may be sufficient for some MFA scenarios. However, there is no differentiation between accessing the account information and access to payments. A service would need to become a TPP with the role Account Information Service Provider (AISP), which would potentially allow asking the user for access to all of his/her account information.

There seems to be no restriction on who might become a TPP [TW Blog]:

> *'All types of companies can become a TPP if they so desire and as long as they fulfil the national requirements for TPP approval. … While the TPP register on a European level will be supervised by the European Banking Authority – EBA – the national financial supervisory authorities are responsible for setting up practices and procedures for TPP approval.'*

Since for GÉANT the national authority would probably be in the Netherlands, a check at the DeNederlandscheBank (DNB) was conducted. They accept TPP applications since February 2019, with just two entities in their AISP registry [DNB] so far. Furthermore, a list of PSD2 use cases for insurance companies [Innopay] includes:

> *'7. Digital identity verification: banks can help in identifying a person during a digital onboarding or digital identity verification process. This functionality is for example already operated by the banks in The Netherlands under the iDIN scheme.'*

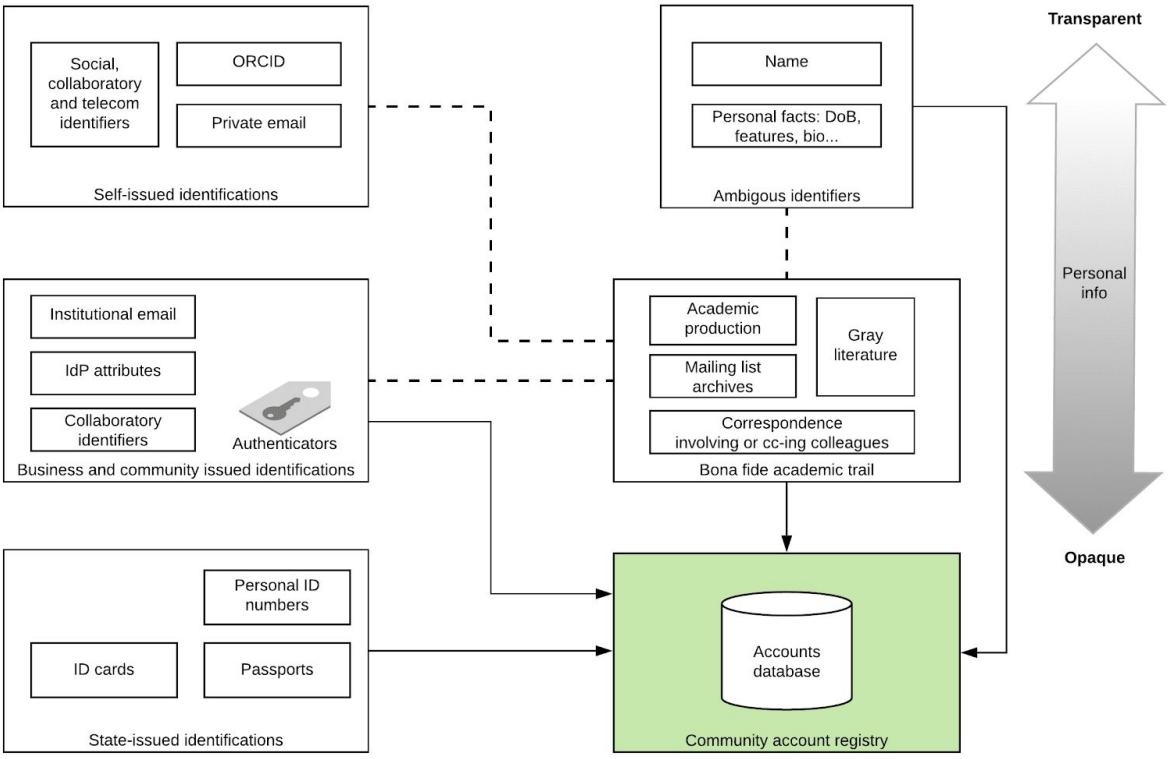
If a service is to include payment of a fee for its use, physical tokens, or other services such as delivery, it ought to assume another TPP role: only a Payment Initiation Service Provider (PISP) can conduct payment transactions.

Re-use of digital trail

Even cheaper is the option to re-use the digital trail of a user. In an academic environment, this may mean the recurring of an institutional email in published papers or the ORCID account. This method yields good results for more academically active participants that are publishing using the same email or ORCID for several years. In their case, by simple email verification, we gain more information than the mere connection to a domain. We get a whole history of activities.

The typical academic research project is organized around a single research subdiscipline. This means that the communities are assembled around a specific research area or topic - like focusing on a single material, e.g. graphene, or a single tool, like a linear particle accelerator.

This means that at least the senior participants of the project will all have a track record of a priori knowledge and activities in the topic in question. When it comes to remote vetting, this knowledge can be exploited to some extent, on top of other methods, to provide additional assurance. This academic trail may include white papers or a dissertation published by the person, using the name, affiliation, email and possibly ORCID identifiers. The produced papers may be listed or accessed by using the subscription-based or open databases and repositories of academic journals, publications and articles. The person may have presented at various conferences before and there may even be video records of that. Also, the person in question may have participated in debates on mailing lists. And most importantly, the person should have gathered substantive knowledge and experience in the research topic.



Serviceable means of personal identification
https://www.lucidchart.com/documents/edit/cec1036e-bc75-4ef9-8509-744a4a20ecb3/0_0

The figure above shows the different aspects of the situation. A person uses a set of identifiers throughout its career and those identifiers end up being bound to the academic career in many ways. For instance, an institutional email or an ORCID account may be used in various impactful publications over several years. These are published by following editorial correspondence or face-to-face meetings at conferences and workshops. Building up a fake identity in a given subdiscipline must take an enormous effort that is premeditated years ahead. For senior participants of a research collaboration, it is typically the case that considerable assurance between 1) there is a strong bond of a person's name to relevant scientific work; 2) there is an affiliation-name relationship that however might change over time; 3) there is a similar name-email relationship (it may also change with the institution); 4) in case ORCID is used extensively, a name-ORCID-subdiscipline relationship may be evident.

Undeniably, face-to-face vetting at a registration desk, conferences or project meetings is the safest methods to facilitate the distribution of second-factor credentials. However, there could be a universe of use cases where good security and assurance is needed (i.e. better than simple self-registration) yet the risks do not justify face-to-face contact. Face-to-face vetting may be the safest bet but also quite expensive.

Relying on the academic trail we may approximate the assurance of the face-to-face method to some extent by operating remotely.

Here is how this kind of process could look like. This process can only be executed by people knowledgeable in the given subdiscipline.

1) An account is created at the target system by the applicant where we eventually want to enable two-factor authentication.

2) The email address ownership is verified by a sending of a simple soft token to the provided address, while the applicant's name is self-registered.

3) The applicant can provide "evidence" in the form of links or bibliographical information about papers or other publicly available correspondence that ties the name to the email and subdiscipline.

4) A live video call is scheduled, where the applicant is asked to present an ID card.

5) During a live video call, the applicant is interviewed and queried about the subdiscipline, essentially like a simple university exam. Other studies (like the SURF report) focus on whether ID cards can be faked in a live video stream, but an examination of the applicant in terms of academic knowledge or, even better, subject of their academic publications by a person acquainted with the domain is almost as effective in a live stream as in face-to-face.

6) During the call, a token is sent to the email to the applicant's email address, or, optionally in a text message to the mobile number. This token needs to be read back during the call. (In some remote vetting cases done by financial institutions, this is the primary way of tying a mobile number to an account).

   So far, we have established a link between an email address, a name, and a subdiscipline. Faking this would require tampering the video feed, the ability to reliably answer discipline-specific questions without hesitation, and in the case when there is a pre-existing document trail, a premeditated effort to build up or hijack a fake identity. Therefore, while this method does not reach the assurance of the face-to-face meeting, it does not fall very short from that.

7) Finally, while the video chat is underway, it is possible to register the second factor, e.g. by using a smartphone application.

   Alternatively, if there is an academic trail that can be established by a document, and the value of the defended resource is not too high, we may get a good identity even without the video chat.

4.4 Comparison of costs

When grading of authentication assurance options for any organisation or collaboration, the associated cost, use of resources and the organisational burden should be considered.

The used assurance scheme may include additive points that support specific subcriteria. The following list indicates the methods that may provide such points. In it, different approaches are classified according to the expected costs:

- **CHEAP:** A website where people can register by email and name
  a) A registered request should be generated for you
  b) The existing person's email and a new password is used
     - Institutional email domain (+points)
  c) The use of eduGAIN IdP (+points)
  d) Opportunistic peer-to-peer piggybacking of regular face-to-face meetings to verify individuals (and at the same time distribute tokens, if needed)
  e) Confirmation of a project participant through corroboration by N confirmed participants (web of trust)

- **REQUIRING EFFORT FOR EACH APPLICANT:** Asynchronous checks
  a) Academic papers with the same name, ORCID and email (+points)
     - Grey literature, email collections (+points)
  b) Digital trail within the actual context (e.g. minutes, previous project deliverables)

- **REQUIRING SYNCHRONISED EFFORT:** Video calls for vetting of ID documents and, to some extent, vetting of knowledge (+points)

- **PER ITEM COSTS:** Piggybacking of services such as postal delivery with the check of recipient's ID or Western Union

- **BOTH FIXED AND TRAVEL COSTS:** Formal face-to-face meetings (this approach provides the highest assurance standard)
  a) Distributed registration desk
  b) Central registration desk

Distribution of physical tokens in a small research community relying on in-person meetings should not be mixed up with distributed identity vetting by dedicated personnel. The postal delivery-based vetting is distributed; however, its cost is rather high, especially if it is also used to deliver physical tokens. The bank transfer-based method does not require any physical exchange, has its forerunners in PayPal and similar services, but the additional security is limited to the vetting of the bank and the uniqueness of the principal's name.

5 Summary of findings

In chapter 1 and 2, this report is scoping the investigation area to typical academic research collaborations in Europe and potentially global. These projects have some commonalities that can be exploited to keep the costs down when it comes to remote vetting. One such feature is that at least a majority of senior participants usually have face-to-face meetings at least at the beginning of the project; the other common feature is the fact that these

collaborations are organized around a sub-discipline, allowing for some sort of examination of knowledge, not only the identity of a prospective new participant, making it harder for a malicious party to impersonate someone.

Chapter 3 provides a Taxonomy of Actions adapted from ITU-T X.1254 entity authentication assurance framework [ITU] terminology to establish a language for discussing the different methods. This allows us to communicate the subtle but important differences between what certain actions achieve. For instance, identity proofing and information verification can be conceptually distinguished from factor binding and activation, even though they may happen during the same process.

Chapter 4 describes several existing vetting methods from industry and academia but also summarises the findings of other research. After considering the cost of the most commonly used vetting methods, like the central registration desk or the distributed face-to-face method, we list some possibly unconventional but potentially useful approaches. These include piggybacking existing parcel delivery or money transfer services including the features provided by PSD2, as well as relying on not only identity documentation but also academic history to bind the user to various associated identifiers.

Finally, our report is concluded by a recommendation section for future research and actions, including interviews, explorations of LoA expressions and case studies.

6 Recommendations for further steps

6.1 Conduct interviews

The next steps may involve carrying out interviews with interested parties to examine whether some of the example flows would apply to a given research and education collaboration. Furthermore, organizations that already have a vetting mechanism in place might be of interest as well to get to know their considerations and reflections but also to share experiences. The following list provides an overview of potential interviewees:

- Authors of the SURF report
- Bank IT
- CERT Authority
- eduid
- BBMRI
- CORBEL
- Elixir
- LS-AAI
- HPC community
- eduTEAM, especially guest IdPs
- SPs related to health, genomics

Several potential exploratory questions have been identified as well:

- Target/eligible audience (to whom this method is available)?
- Approximate size of the eligible or covered population and its overlap with the audience with federated NREN/GN identities?
- Applicability for the target purpose and scope?
- Level of identity assurance?
- Any policies or documents about issuance/revocation procedures?
- Does it provide or supports some form of authenticity check?
- Any machine-readable issuance/revocation lists?
- Remote, centralised or distributed?
- Current vetting capacity?
- Estimated cost per user?
- Estimated fixed or periodic costs?
- Does the vetting process provide the user with something that could be used for MFA? Ways to piggyback the two?
- What is good with this solution?
- What is bad about it?
- Are there some interesting technical details or episodes you would like to share?
- Any technical specs/docs?

6.2 Explore Level of Assurance (LoA) frameworks

Assurance levels characterise the degree of confidence in electronic identification means in establishing the identity of a person, thus assuring that the person claiming a particular identity is, in fact, the person to which that identity was assigned.

As discussed in the previous sections, there are many available vetting mechanisms. Whereas in-person vetting methods, like a central registration desk, offer a pretty high degree of assurance, mechanisms like merely relying on a verified email address offer a rather weak assurance. Hence, before implementing a practical identity vetting and token registration prototype, Level of Assurance (LoA) frameworks could be a useful baseline for future elaborations of vetting schemes.

There are several LoA frameworks, some of them being listed here:

- **NIST Special Publication 800-63-3** [NIST] is a comprehensive assurance framework including three categories: Identity Assurance Levels (IAL), Authenticator Assurance Levels (AAL) and Federation Assurance Levels (FAL). IAL defines strict requirements, making it hard to implement them in our scope.

- **Kantara Identity Assurance Framework** (Kantara Classic) [Kantara] is based on NIST 800-63-2. Allows some flexibility due to the definition of high-level requirements.

- **IGTF Profiles of Authentication Assurance** [IGTF] originates from the grid infrastructure and consists of four technology-agnostic authentication assurance levels.

- **eIDAS Regulation** [eIDAS Regulation, eIDAS Implementing Acts] is an EU regulation that sets out rules for electronic identification and trust services. As identification is an element of national sovereignty, it cannot be regulated on the EU level. eIDAS, therefore, uses a lightweight approach of mutual recognition of identification schemes and assurance levels. The Member States have freedom on how to trust the electronic identification means issued in other members. This is facilitated with the eIDAS notification process. It also outlines assurance levels **low**, **substantial** and **high**, but without any particular meaning except for their relationships within and across counties in terms of their practical (downward) substitutability on the base of each individual LoA scheme. eIDAS refers to ISO/IEC 15408 "Information technology – Security techniques – Evaluation criteria for IT security" and ISO/IEC 18045 "Information technology – Security techniques – Methodology for IT security evaluation" as the starting points for the elaboration of such schemes. These standards are freely available at [Common Criteria], with CCPART1-3 being equivalent to ISO/IEC 15408 and CEM to ISO/IEC 18045. Instead of defining the three levels, eIDAS refers to the Large-Scale Pilot STORK and ISO 29115 levels 2, 3 and 4 as the base for minimum technical requirements, standards and procedures for the mentioned three levels. Also, the requirements should be technology-neutral, making it possible to achieve them through different technologies.

- **REFEDS Assurance Framework** [RAF] provides - amongst other criteria - three levels of "*Identity proofing and credential issuance, renewal and replacement*" requirements, with levels related to Kantara, IGTF and eIDAS:

    - ***Low*** *(self-asserted)*
    - ***Medium*** *(e.g. postal delivery)*
    - ***High*** *(e.g. face-to-face with the check of ID documents)*

When relating peer-to-peer and web of trust schemes to LoAs and upon looking at the eIDAS guidance for the application of the levels of assurance [eIDAS LoA Guidance, eIDAS Regulation] it seems more reasonable to expect the need for fixed algorithmic criteria for levels than more abstract and additive points/credits. However, these points could be limited to specific appropriate subcriteria within a larger fixed scheme. For example, in academic research projects, confirming participants in person by trusted mediators (akin to RA), or corroboration by a proscribed number of confirmed participants could be used to avoid physical meetings for linking credentials with participants. In addition, the abolishment of physical meetings requires the use of digital factors instead of physical tokens, which should be also attested by corroborators. This is an example of a restricted application of the decentralized trust model used by the PGP web of trust. This may be further extended by

allowing transitive indirect trust, depending on diminishing points after every hop (weighted graph edges), or tools and metrics that identify and quantify strongly connected sets of individuals, e.g by mean shortest distances.

6.3 Prototype practical scenarios

We have created some example flows that illustrate the usage of the concepts and means described in this document. A simplified BPMN flow appropriate for research projects is given in the introduction. More specified flows, such as those with video calls, dedicated mobile applications or registration desks are provided in the Appendix. These flows can be used as a starting point or, for example, for elaborating custom enrolment processes suitable for specific communities, user subgroups (for which some alternative paths may be necessary), organisational arrangements and applied technologies and tools.

The MFA vetting should be prototyped with at least one interested party, most likely by piloting support for token distribution and management for academic research projects [GN4-3 WP5 Projects Vetting Flow].

6.4 Test, elaborate and integrate novel technologies

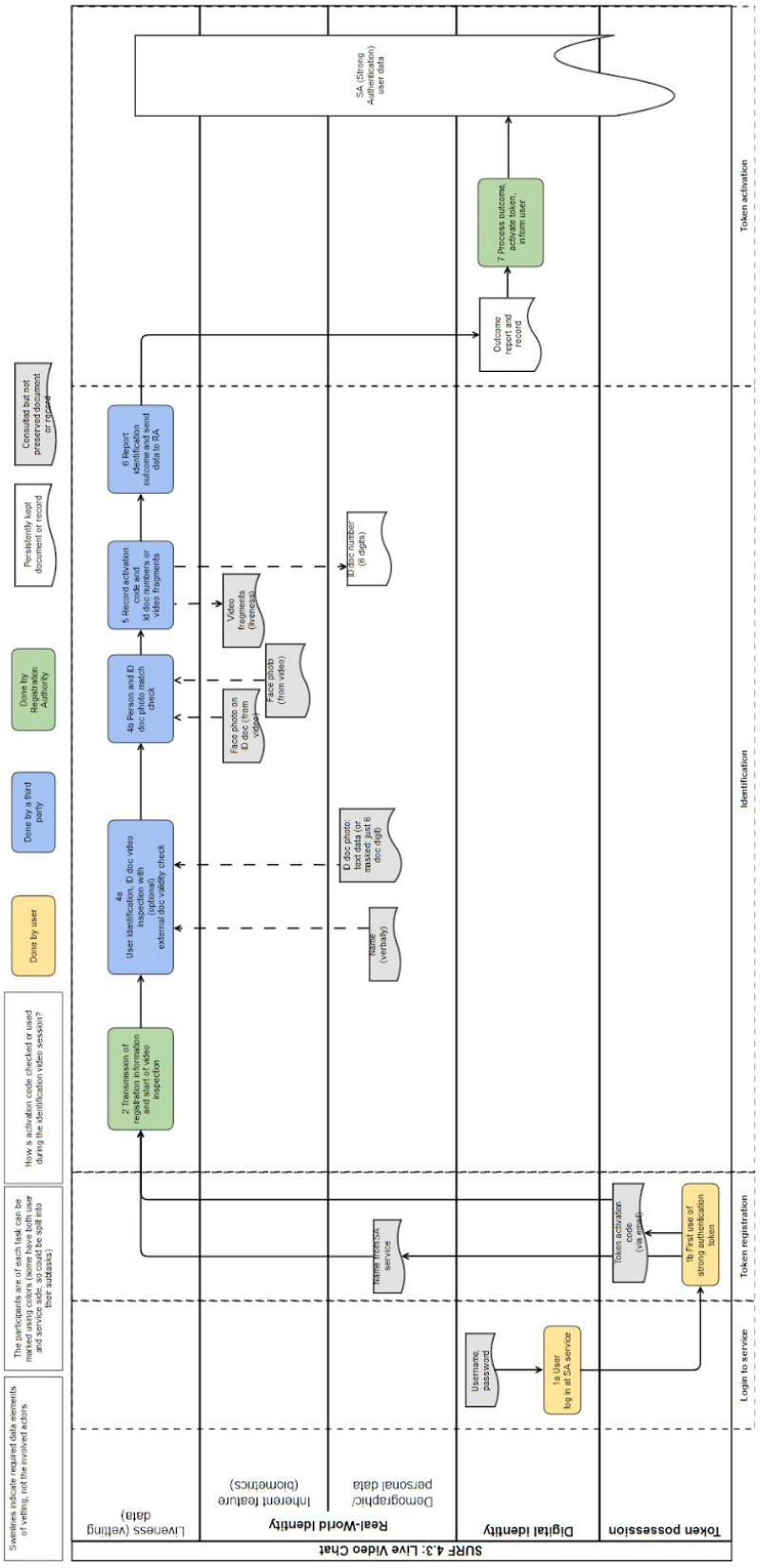
New enrolment flows can be designed and the established ones modified, optimised, or partially automated by using the emergent and typically software-based products and services. These tools could assist in checking the proofs, liveness or identity. Many are described in this report, but the most viable examples worth investigating are:

- Mobile applications to access biometrics, check liveness and read and verify RFID/NFC and biometrics enabled ICAO-compliant passports and other personal documents.
- PSD2 to access bank account or money transfers data to verify information provided by applicants.
- Integration with issuers of various identity documents and record who or whose products can attest applicants' identity.

Appendix: Detailed BPMN workflows

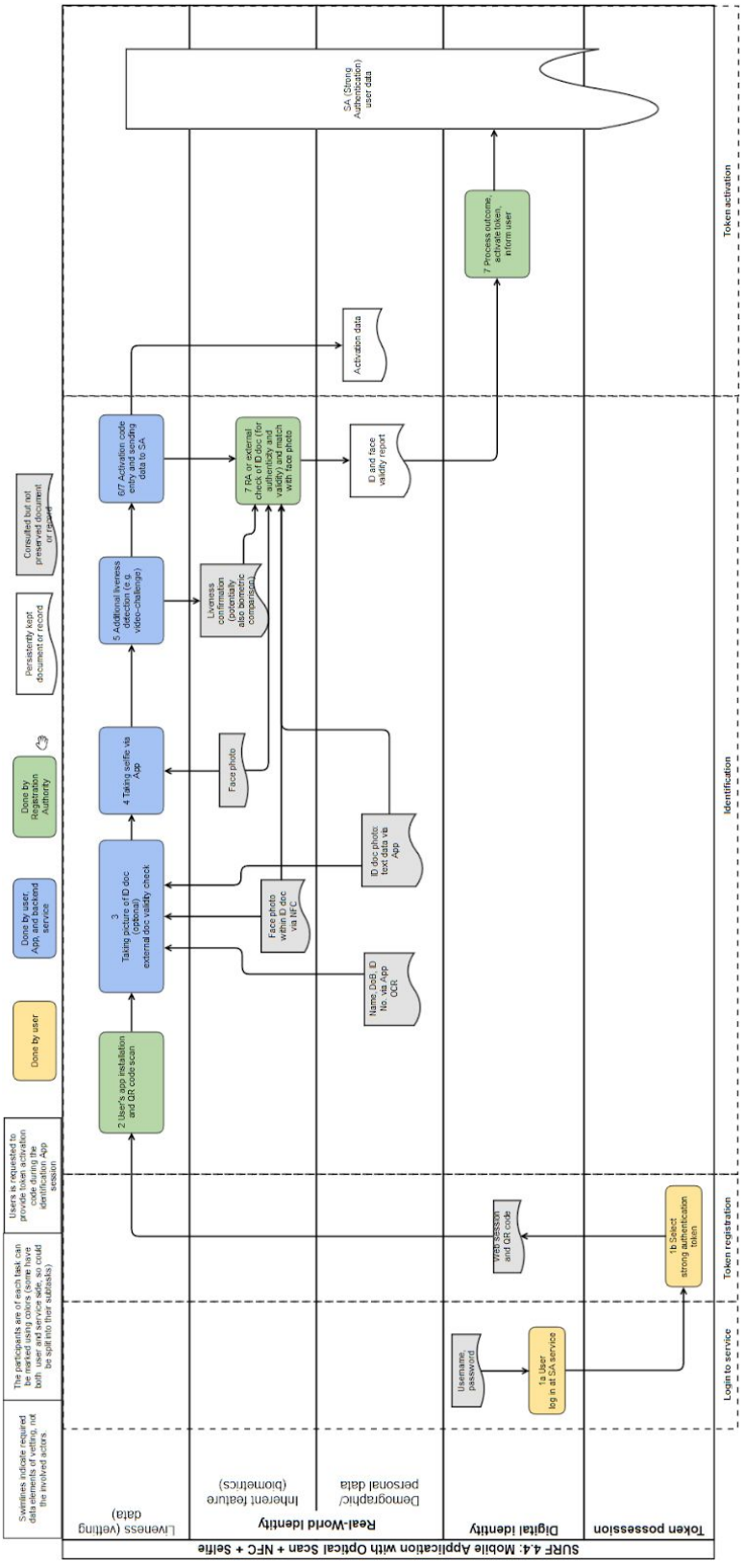
SURF document 4.3: Live video call

https://wiki.geant.org/display/gn43wp5/Flow%3A+SURF+4.3+Live+Video+Chat



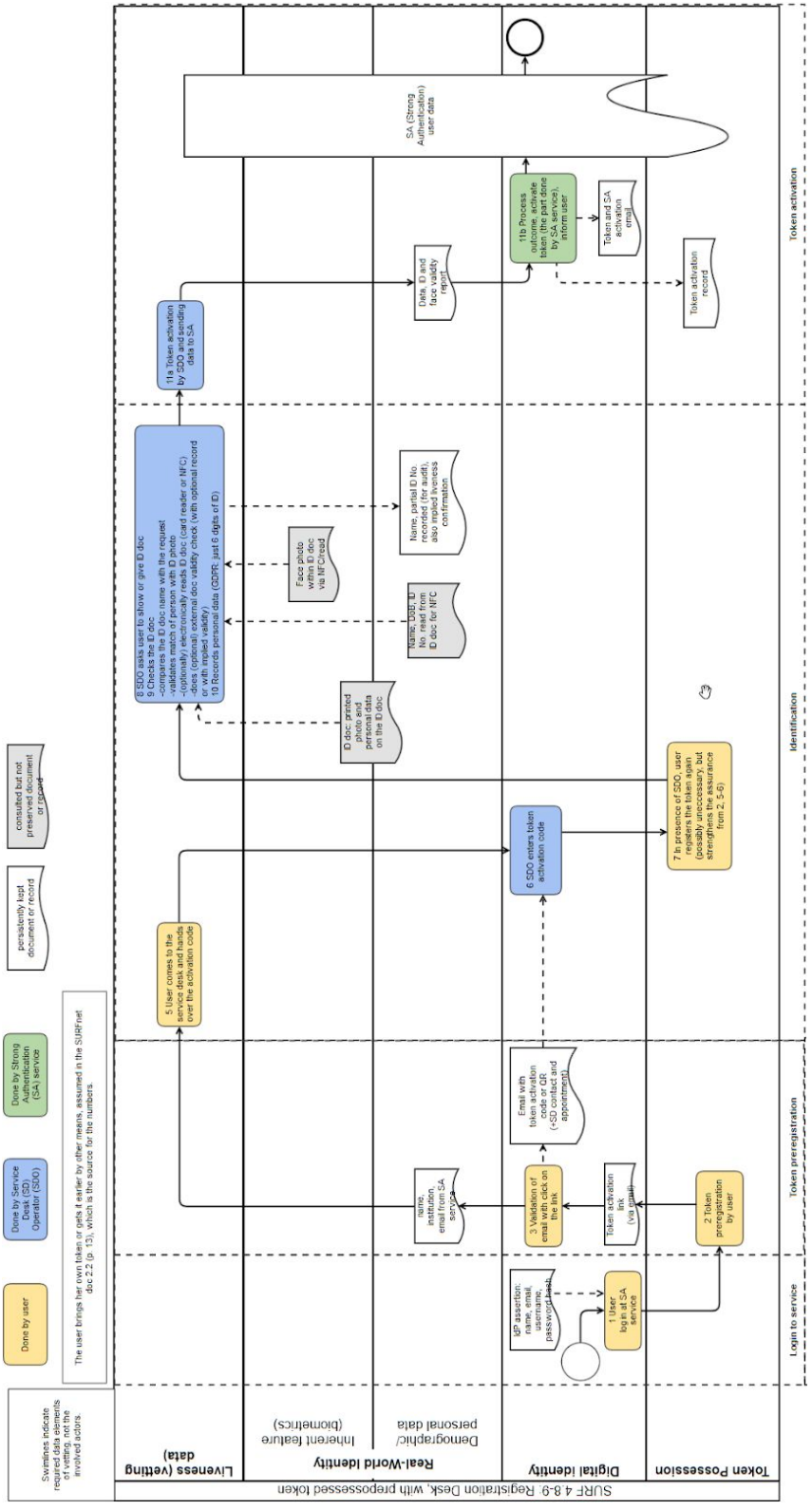
SURF document 4.4: Mobile application with optical scan +NFC +Selfie

https://wiki.geant.org/pages/viewpage.action?pageId=123778436



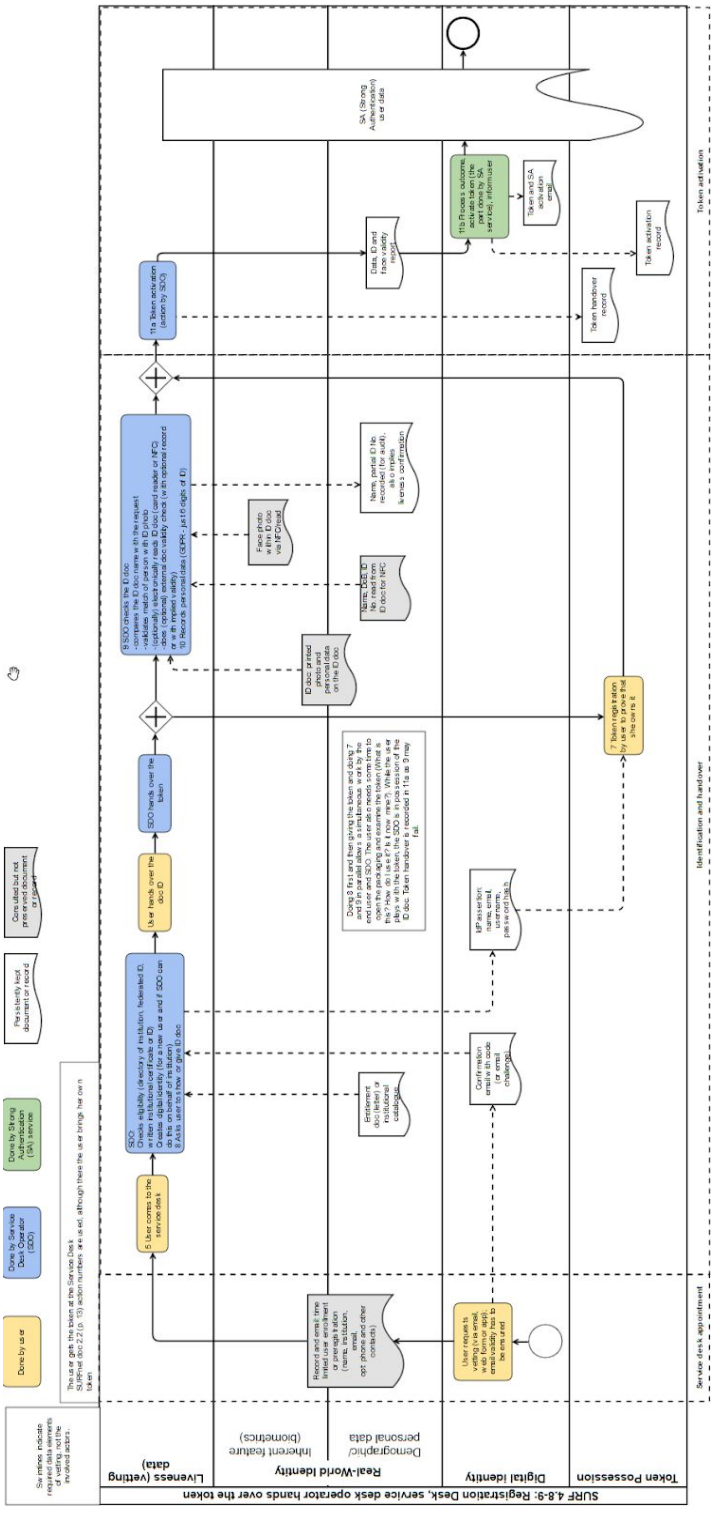
SURF document 4.8-9: Registration desk, prepossessed token

https://wiki.geant.org/display/gn43wp5/Flow%3A+SURF+4.8-9+Registration+Desk%2C+with+prepossesed+token



SURF document 4.8-9: Registration desk, operator hands over the token

https://wiki.geant.org/display/gn43wp5/Flow%3A+SURF+4.8-9+Registration+Desk%2C+service+desk+operator+hands+over+the+token



References

Berlin Group	https://www.berlin-group.org/nextgenpsd2-downloads
DNB	https://www.dnb.nl/en/supervision/public-register/WFTRI/index.jsp
Common Criteria	www.commoncriteriaportal.org/cc
eIDAS Regulation	https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2014.257.01.0073.01.ENG
eIDAS Implementing Acts	https://ec.europa.eu/futurium/en/content/eidas-implementing-acts
eIDAS LoA Guidance	https://ec.europa.eu/cefdigital/wiki/download/attachments/40044784/Guidance%20on%20Levels%20of%20Assurance.docx
GN4-3 WP5 Projects Vetting Flow	https://wiki.geant.org/display/gn43wp5/Flow%3A+Vetting+of+Research+Projects+Participants
idin	https://www.idin.nl/en/ in The Netherlands
IGTF	https://www.igtf.net/ap/authn-assurance/
Innopay	https://www.innopay.com/en/publications/insurance-and-open-banking-wave-seven-use-cases
ITU	https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-X.1254-201209-I!!PDF-E&type=items
Kantara	https://kantarainitiative.org/trustoperations/kantara-classic/
NIST	https://pages.nist.gov/800-63-3/
RAF	https://refeds.org/assurance
SURF Report	https://www.surf.nl/files/2019-02/report%20remote%20vetting%20for%20surfconext%20strong%20authentication.pdf
TW Blog	https://equensworldline.com/en/home/blog/2018/april-18/20180410-five-important-steps-to-become-a-third-party-provider-following-psd2.html