Securing Distributed Research

Hannah Short
CERN IT-DI-CSO

Securing Distributed Research

Identity Federation for Research

Global science calls for global infrastructure. A typical large-scale research group will use a suite of international services and involve hundreds of collaborating institutes and users from around the world. How can these users access those services securely? How can their digital identities be established, verified and maintained?

We will explore the motivation for distributed authentication and the ways in which research communities are addressing the challenges. We will discuss security incident response in distributed environments - a particular challenge for the operators of these infrastructures. Through this course you should gain an overview of federated identity technologies and protocols, including certificates, SAML and OAuth2.

Who am I?

- Member of CERN’s IT Department
- Working on Trust and Identity for CERN and WLCG
- hannah.short@cern.ch

What are we talking about?

- Authentication, Authorisation and Identity
- Authentication & Authorisation for Distributed Communities
- Security Incident Response

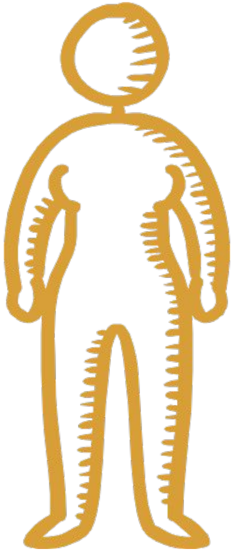
What am I hoping that you will remember?

- The global research community is increasingly connected through shared use of digital identities
- This brings benefits and also challenges
- There are multiple ways of doing it

If you are developing a service that needs authentication or authorisation, come back and look at this!

Authentication, Authorisation and Identity

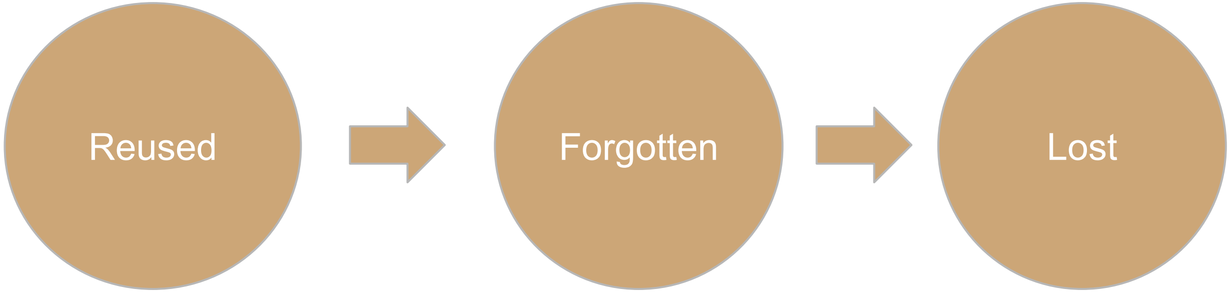
Authentication, Authorisation and Identity



Authentication, Authorisation and Identity

Traditional Online Identity

- Bits of identity scattered through the web
- Very different idea of “me”
- 100s of username and password pairs



Reused

Forgotten

Lost

Authentication, Authorisation and Identity

The Evolving Online Identity

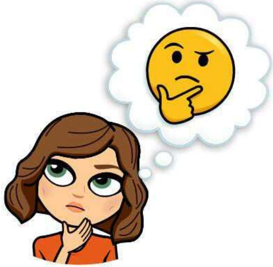
- Credibility is key
- Use of a single, unified identity is becoming increasingly possible
  - Link accounts
  - Log in with Social ID
- Being able to grant permissions to well defined identity attributes or capabilities is gaining importance

Donald J. Trump ✓

@realDonaldTrump

#MakeAmericaGreatAgain #Trump2016 #Facebook: facebook.com/DonaldTrump #Instagram: instagram.com/realdonaldtrum...

What impact does this have for Research?



Why do we need Authentication & Authorisation?

- Confidentiality

The final research may be public but, until that point confidentiality matters!

For example, there is a deliberate separation between certain experiments.

Why do we need Authentication & Authorisation?

- Confidentiality
- Traceability

When something goes wrong, we need to be able to trace back to the user.

For example, a physicist submits a job that seg faults. If we can work out where the job came from we can get in contact and help.

Why do we need Authentication & Authorisation?

- Confidentiality
- Traceability
- Attribution

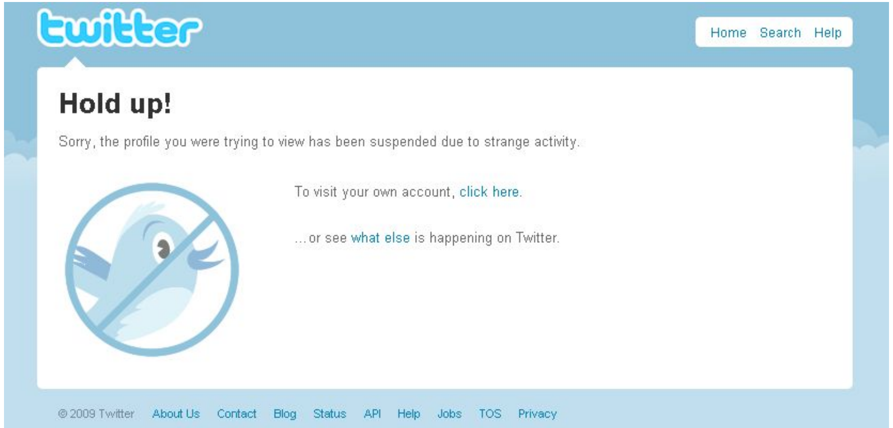
Having a stable, reliable identifier allows research to be properly attributed. Identity changes present problems - e.g. standard advice is not to change your name when you get married.

ORCID provides life-long identifiers to researchers, to attach to publications, grant requests etc

Visit *https://blog.inspirehep.net/2015/04/what-is-orcid-and-how-can-it-help-you/*
and https://home.cern/cern-people/updates/2018/01/get-yourself-orcid

Why do we need Authentication & Authorisation?

- Confidentiality
- Traceability
- Attribution
- Suspension



Hold up!

Sorry, the profile you were trying to view has been suspended due to strange activity.

To visit your own account, click here.

...or see what else is happening on Twitter.

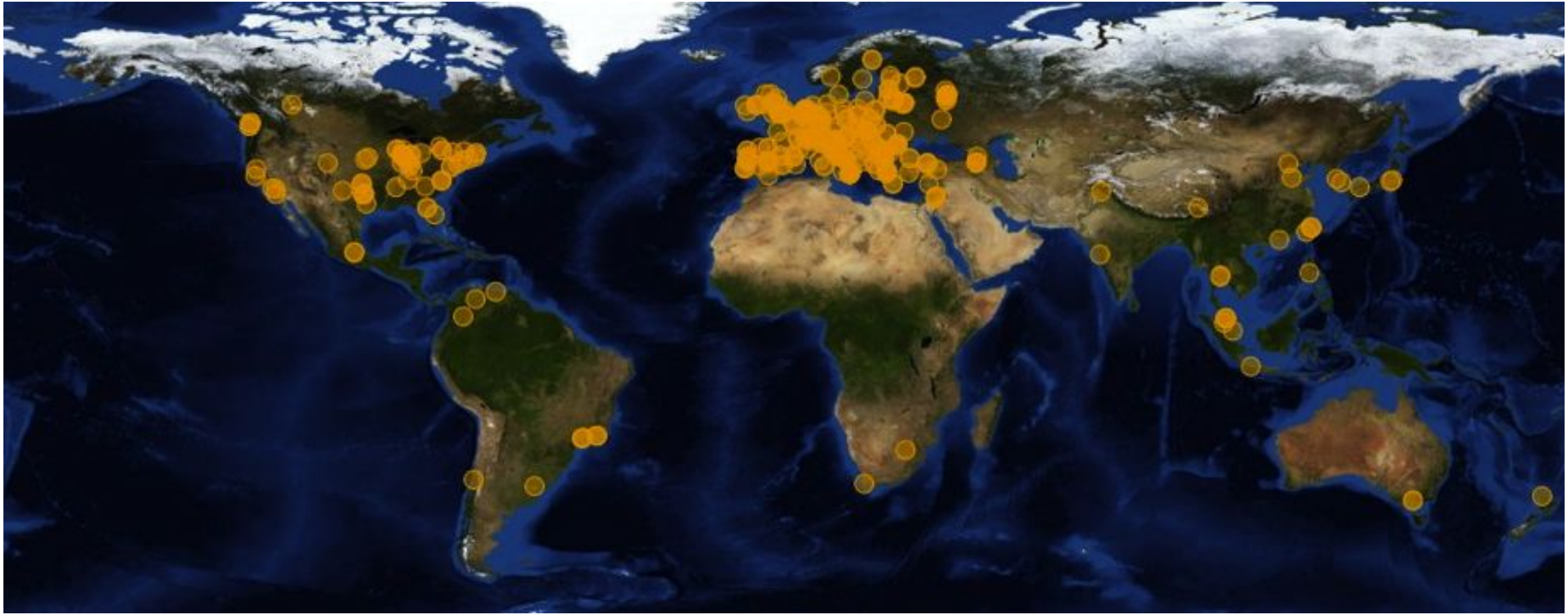
© 2009 Twitter About Us Contact Blog Status API Help Jobs TOS Privacy

What if a user's identity is compromised? Can we isolate the identity and suspend it?

Otherwise, do we have to stop the jobs from an entire experiment?

Authentication & Authorisation for Distributed Communities

Authentication for Distributed Communities

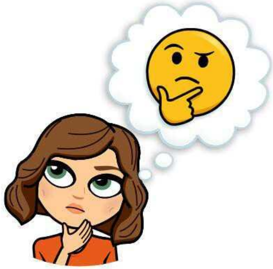


Authentication for Distributed Communities

The problem

- Large, global user community
- Working on the same infrastructure
- Don’t necessarily know each other
- Don’t necessarily ever meet

How can we securely provision digital identities?



Authentication for Distributed Communities

Who knows the users?

- The Laboratory?
- The Infrastructure?
- The Experimental Group?
- The Home Organisations?
- A trusted 3rd party?

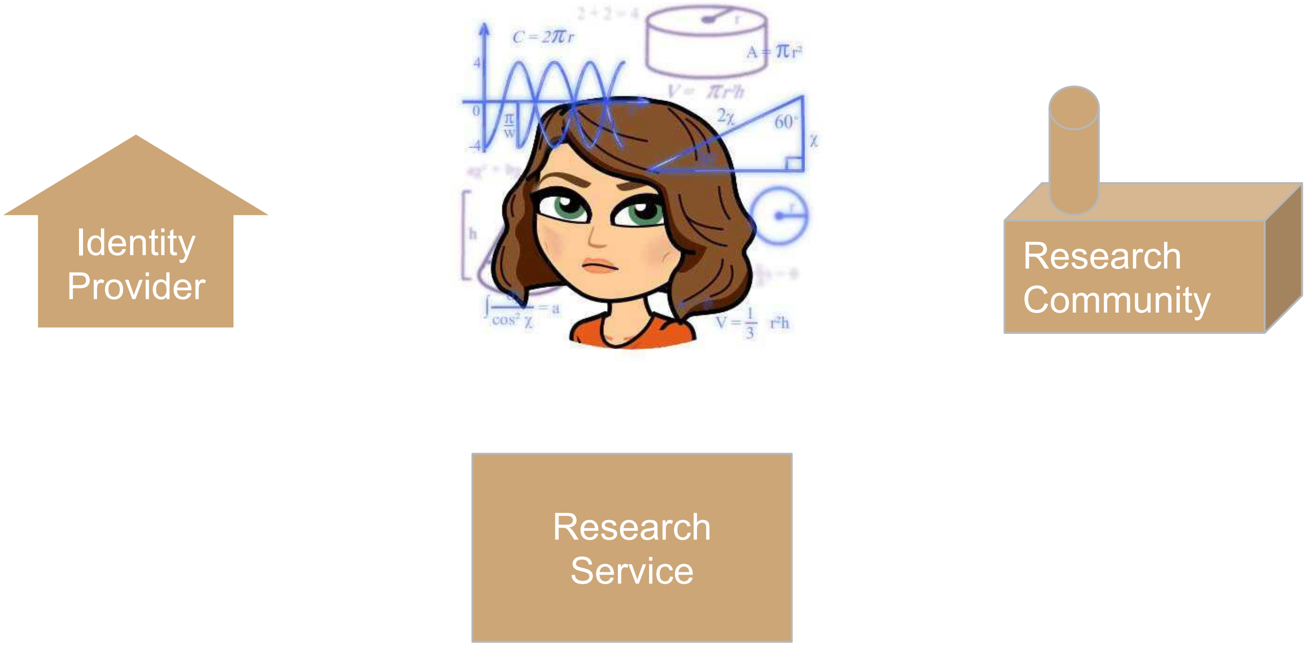
Typically, the **home organisation** may have the most current information. Or, potentially, a trusted 3rd party like a government, bank or organisation specialised in identity vetting

Authorisation for Distributed Communities

By contrast, **Experimental Groups or Research Communities** may be better placed to know

- Which group you belong to
- Which roles you should have, e.g. user, admin, super-user
- When you need to accept new or changed policies, e.g. Acceptable Use Policy (= “no bitcoin”)

Putting the Pieces Together

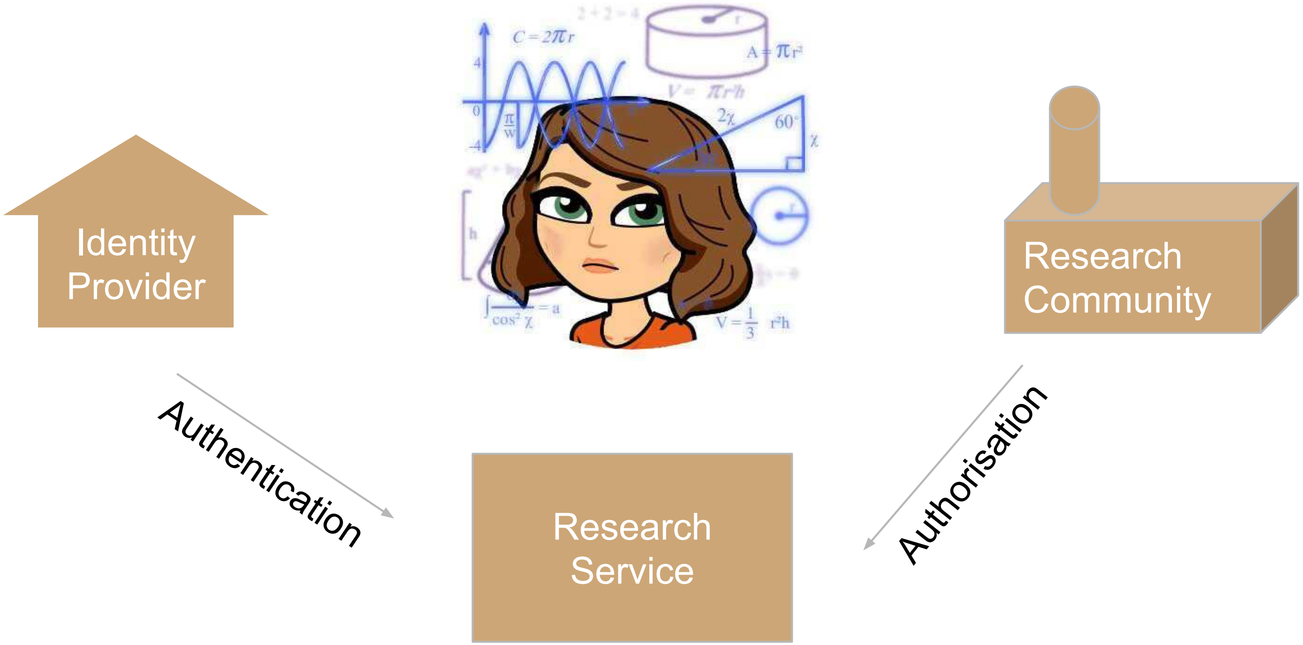


Identity Provider

Research Community

Research Service

Putting the Pieces Together



Identity Provider

Research Community

Authentication

Research Service

Authorisation

How does it work?

How does it work?

There are multiple possibilities, we’ll focus on the three most widely used methods for distributed authentication

- Certificates
- SAML
    - XML bundles
- OAuth2
    - Tokens

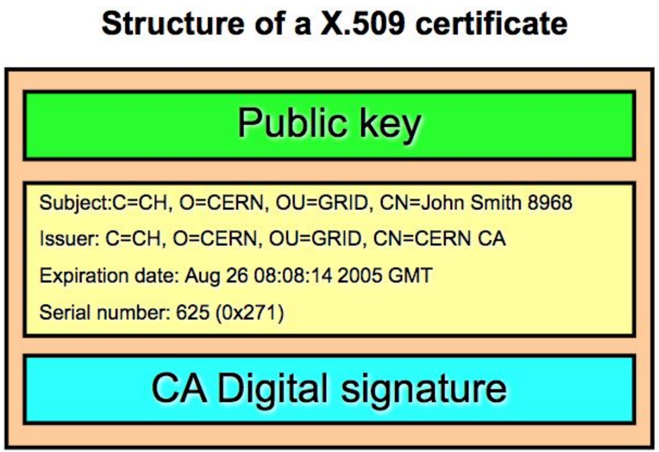
Certificates

“In cryptography, a public key certificate, also known as a digital certificate or identity certificate, is an electronic document used to prove the ownership of a public key.” Wikipedia

Certificates

Certificates can be granted to any entity - service, machine or individual

- Digital Identity
- Signed by a Certificate Authority
  - Hash of the certificate, encrypted with the CA's private key
- Typically long lived ~1yr
- Accompanying private key has a password



Structure of a X.509 certificate

Public key

Subject:C=CH, O=CERN, OU=GRID, CN=John Smith 8968

Issuer: C=CH, O=CERN, OU=GRID, CN=CERN CA

Expiration date: Aug 26 08:08:14 2005 GMT

Serial number: 625 (0x271)

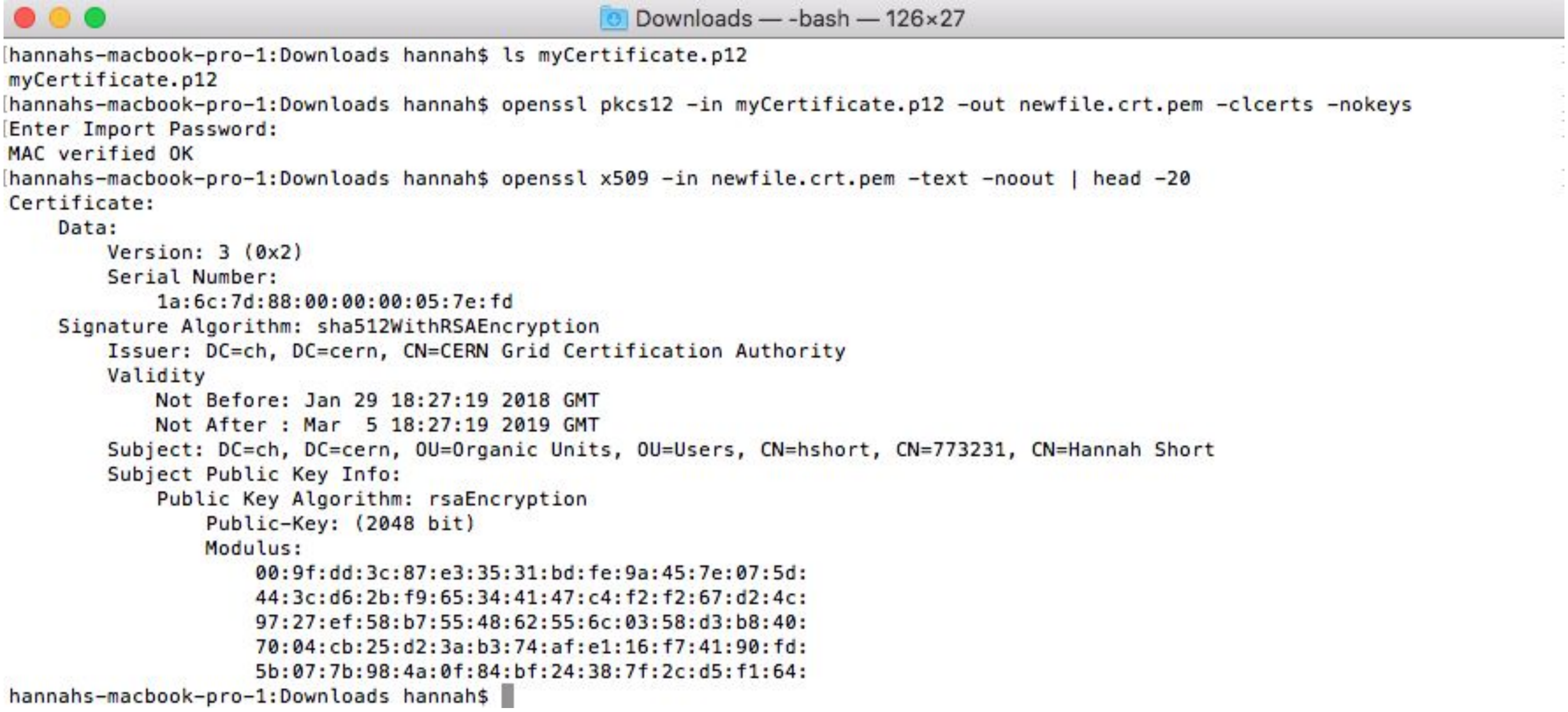
CA Digital signature

http://slideplayer.com/slide/10176602/

Certificates for Research

What does the IGTF do?

- Certificate Authorities regulated by the Interoperable Global Trust Federation (IGTF)
  - Signed by CA **IF** they can validate the identity
  - X509 is the form of certificate used in the Grid
- Authentication = Certificates
- Authorisation = Certificate Extensions

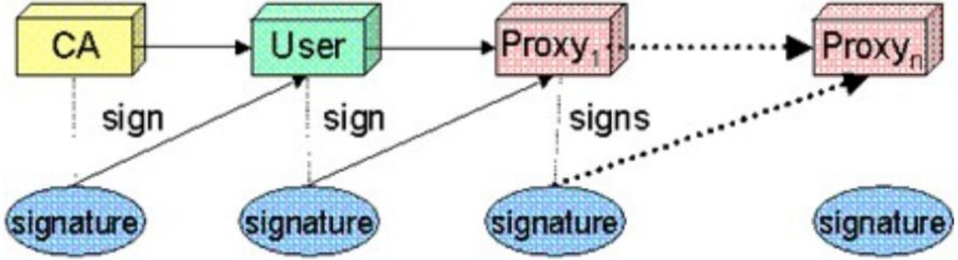


```
[hannahs-macbook-pro-1:Downloads hannah$ ls myCertificate.p12
myCertificate.p12
[hannahs-macbook-pro-1:Downloads hannah$ openssl pkcs12 -in myCertificate.p12 -out newfile.crt.pem -clcerts -nokeys
[Enter Import Password:
MAC verified OK
[hannahs-macbook-pro-1:Downloads hannah$ openssl x509 -in newfile.crt.pem -text -noout | head -20
Certificate:
    Data:
        Version: 3 (0x2)
        Serial Number:
            1a:6c:7d:88:00:00:00:05:7e:fd
    Signature Algorithm: sha512WithRSAEncryption
        Issuer: DC=ch, DC=cern, CN=CERN Grid Certification Authority
        Validity
            Not Before: Jan 29 18:27:19 2018 GMT
            Not After : Mar  5 18:27:19 2019 GMT
        Subject: DC=ch, DC=cern, OU=Organic Units, OU=Users, CN=hshort, CN=773231, CN=Hannah Short
        Subject Public Key Info:
            Public Key Algorithm: rsaEncryption
                Public-Key: (2048 bit)
                Modulus:
                    00:9f:dd:3c:87:e3:35:31:bd:fe:9a:45:7e:07:5d:
                    44:3c:d6:2b:f9:65:34:41:47:c4:f2:f2:67:d2:4c:
                    97:27:ef:58:b7:55:48:62:55:6c:03:58:d3:b8:40:
                    70:04:cb:25:d2:3a:b3:74:af:e1:16:f7:41:90:fd:
                    5b:07:7b:98:4a:0f:84:bf:24:38:7f:2c:d5:f1:64:
hannahs-macbook-pro-1:Downloads hannah$ 
```

Proxy Certificates

Why do we need a proxy? Why not send the certificate?

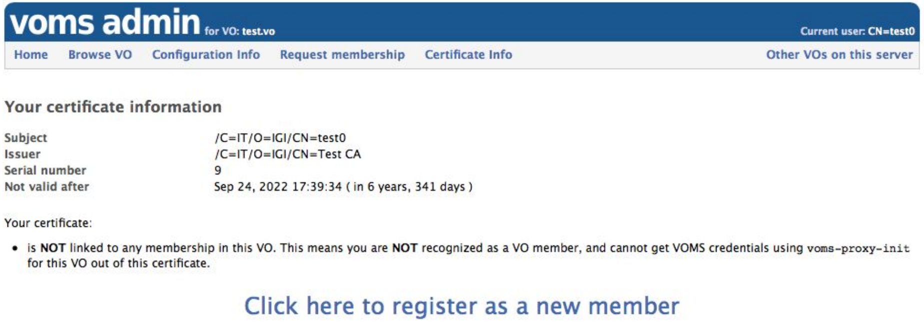
- The user certificate is used to generate and sign a
  - Proxy Certificate
    - Identity of the user
    - Short lived
    - Expiration time
  - Private key
    - No password
    - Readable only by the user
- Proxy and its private key are sent off together and can generate new proxies



http://slideplayer.com/slide/10176602/

VOMS Proxy Certificates

What is a Virtual Organisation?



Your certificate information

Subject	/C=IT/O=IGI/CN=test0
Issuer	/C=IT/O=IGI/CN=Test CA
Serial number	9
Not valid after	Sep 24, 2022 17:39:34 (in 6 years, 341 days)

Your certificate:

- is **NOT** linked to any membership in this VO. This means you are **NOT** recognized as a VO member, and cannot get VOMS credentials using `voms-proxy-init` for this VO out of this certificate.

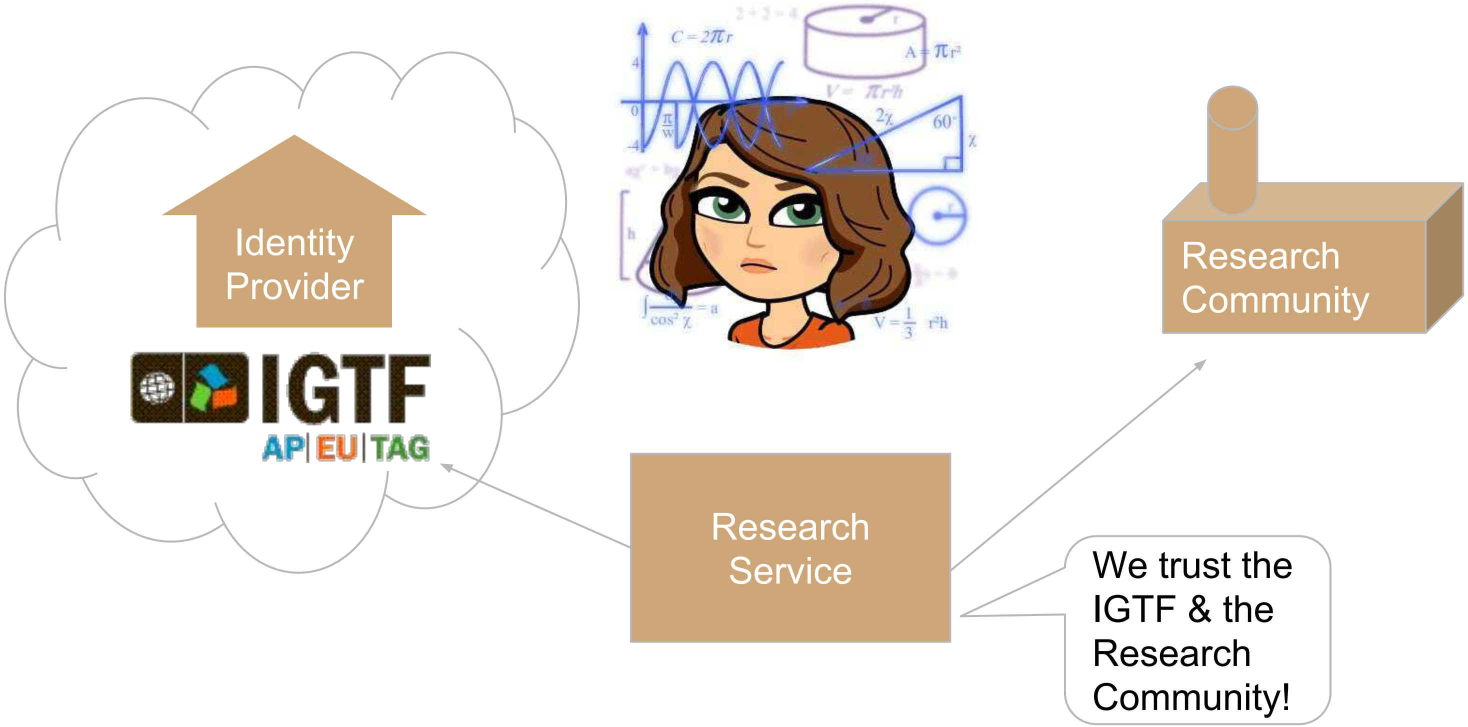
Click here to register as a new member

- **Grid Proxy** = Short lived certificate to be used for authentication to grid services
- **VOMS Extension** = Virtual Organisation specific information, e.g. role and capability
- **VOMS Proxy = Grid Proxy + VOMS Extension**

https://eu-egee-org.web.cern.ch/eu-egee-org/fileadmin/documents/UseCases/ProxyCerts.html

http://toolkit.globus.org/toolkit/docs/5.0/5.0.2/security/gsic/user/ https://www.ietf.org/rfc/rfc3820.txt

Where’s the trust?



Identity Provider

IGTF

AP|EU|TAG

Research Community

Research Service

We trust the IGTF & the Research Community!

Certificates

Good Bits

- Well established technology, services are set up to accept certificates
- Same credential valid for web and non-web

Bad Bits

- Security impact if compromised (and frequently compromised)
- Not user friendly
- Mobility issues

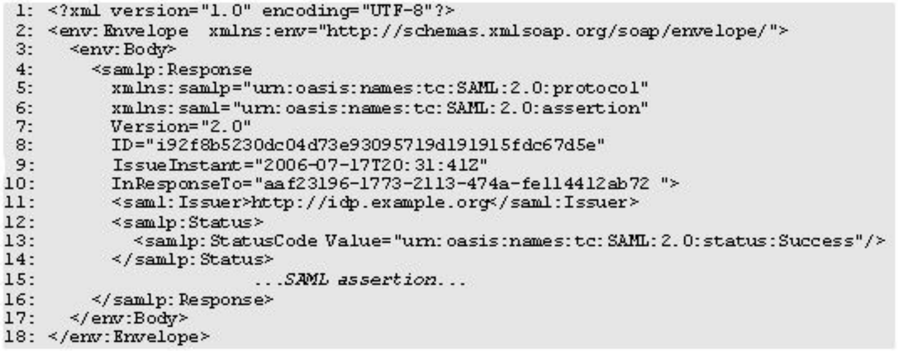
SAML

Security Assertion Markup Language

“Security Assertion Markup Language (SAML, pronounced sam-el[1]) is an open standard for exchanging authentication and authorization data between parties” Wikipedia

SAML

- Often used for Single-Sign-On implementations
- Typically used by the Research and Education sector
- Limited to web services
- Authentication assertions sent as XML packets
    - Can be encrypted or not
    - Contain user attributes



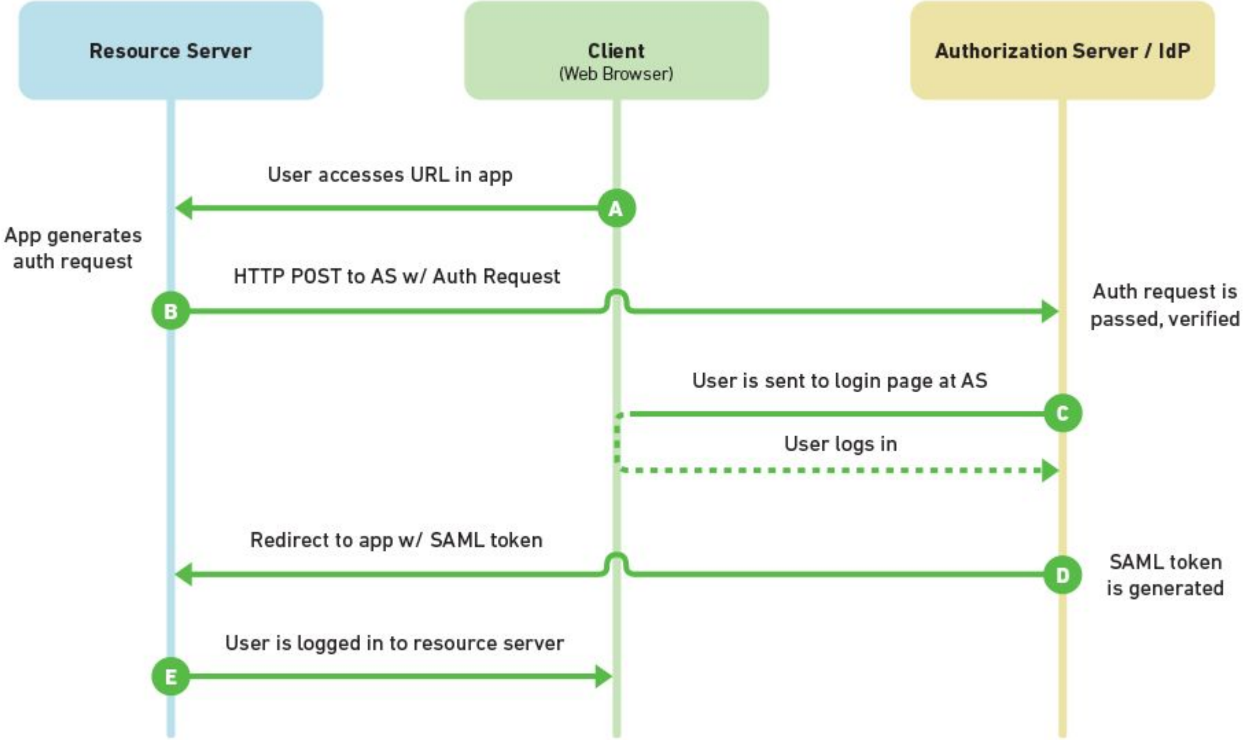
```xml
<?xml version="1.0" encoding="UTF-8"?>
<env:Envelope  xmlns:env="http://schemas.xmlsoap.org/soap/envelope/">
  <env:Body>
    <samlp:Response
      xmlns:samlp="urn:oasis:names:tc:SAML:2.0:protocol"
      xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion"
      Version="2.0"
      ID="i92f8b5230dc04d73e93095719d191915fdc67d5e"
      IssueInstant="2006-07-17T20:31:41Z"
      InResponseTo="aaf23196-1773-2113-474a-fe114412ab72 ">
      <saml:Issuer>http://idp.example.org</saml:Issuer>
      <samlp:Status>
        <samlp:StatusCode Value="urn:oasis:names:tc:SAML:2.0:status:Success"/>
      </samlp:Status>
                    ...SAML assertion...
    </samlp:Response>
  </env:Body>
</env:Envelope>
```

Figure 10: Response in SOAP Envelope

http://docs.oasis-open.org/security/saml/Post2.0/sstc-saml-tech-overview-2.0.html

SAML Protocol

- Client
  - User on their browser
- Resource Server
  - A website requiring authentication
- Authorization server/IdP
  - Home Organisation


Resource Server
Client
(Web Browser)
Authorization Server / IdP
User accesses URL in app
A
App generates auth request
HTTP POST to AS w/ Auth Request
B
Auth request is passed, verified
User is sent to login page at AS
C
User logs in
Redirect to app w/ SAML token
D
SAML token is generated
User is logged in to resource server
E

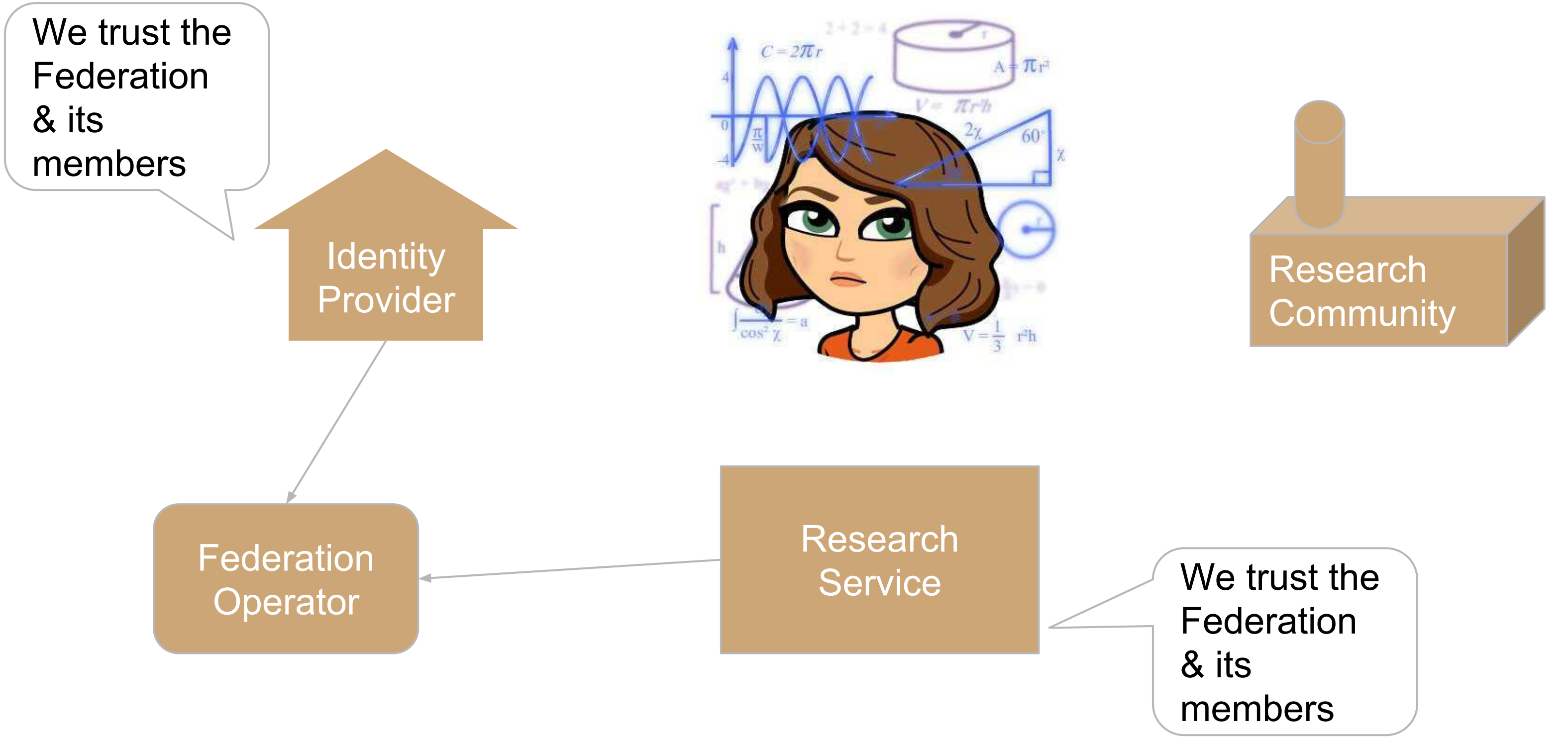
https://www.mutuallyhuman.com/blog/2013/05/09/choosing-an-sso-strategy-saml-vs-oauth2/

SAML Trust Federations

A group of Service Providers and Identity Providers that have agreed to work together.

- Federation metadata collects XML descriptions of each organisation, along with their certificate
- Federation metadata is signed by the Federation and distributed to all members
- Everyone has access to everyone’s certificates, issued by a trusted source

Where’s the trust?



We trust the Federation & its members

Identity Provider

Federation Operator

Research Service

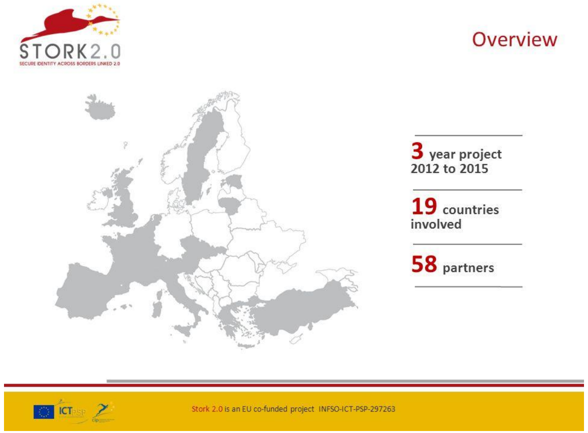
Research Community

We trust the Federation & its members

Interfederation Examples

STORK

“STORK project makes it possible for millions of EU citizens who are resident in a Member State other than their own or work in one country and live in another one to access online public services wherever they are located.”



STORK2.0

SECURE IDENTITY ACROSS BORDERS LINKED 2.0

Overview

3 year project
2012 to 2015

19 countries
involved

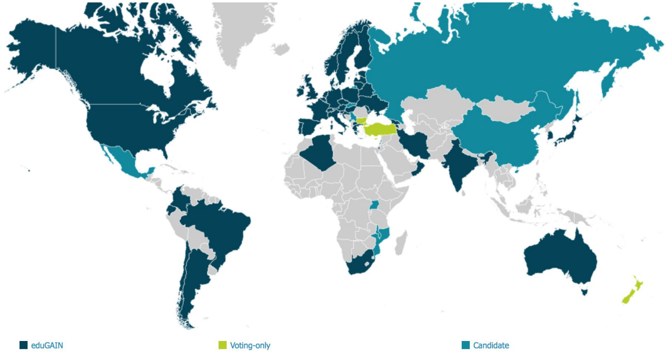
58 partners

Stork 2.0 is an EU co-funded project INFSO-ICT-PSP-297263

http://slideplayer.com/slide/10363474/

eduGAIN

“The eduGAIN service interconnects identity federations around the world, simplifying access to content, services and resources for the global research and education community. eduGAIN enables the trustworthy exchange of information related to identity, authentication and authorisation (AAI). ”



https://technical.edugain.org/

SAML

Good Bits

- Mature, scalable federations
- Secure protocol

Bad Bits

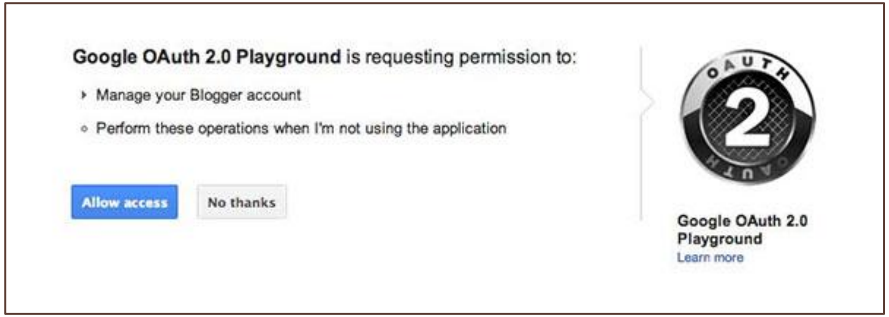
- Only works for web services
- Significant implementation effort

OAuth2

“OAuth is an open standard for access delegation, commonly used as a way for Internet users to grant websites or applications access to their information on other websites but without giving them the passwords” Wikipedia

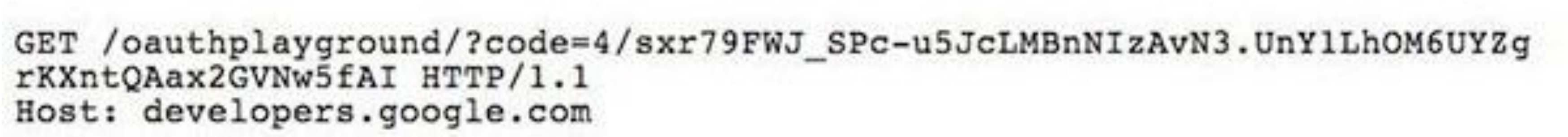
OAuth2

- Typically used by Social Providers, i.e. “GAFA” (Google, Apple, Facebook, Amazon)
- Relies on bearer tokens, i.e. opaque strings signed by the “authorisation service”
- Non-web and API friendly



Google OAuth 2.0 Playground is requesting permission to:

- Manage your Blogger account
- Perform these operations when I'm not using the application

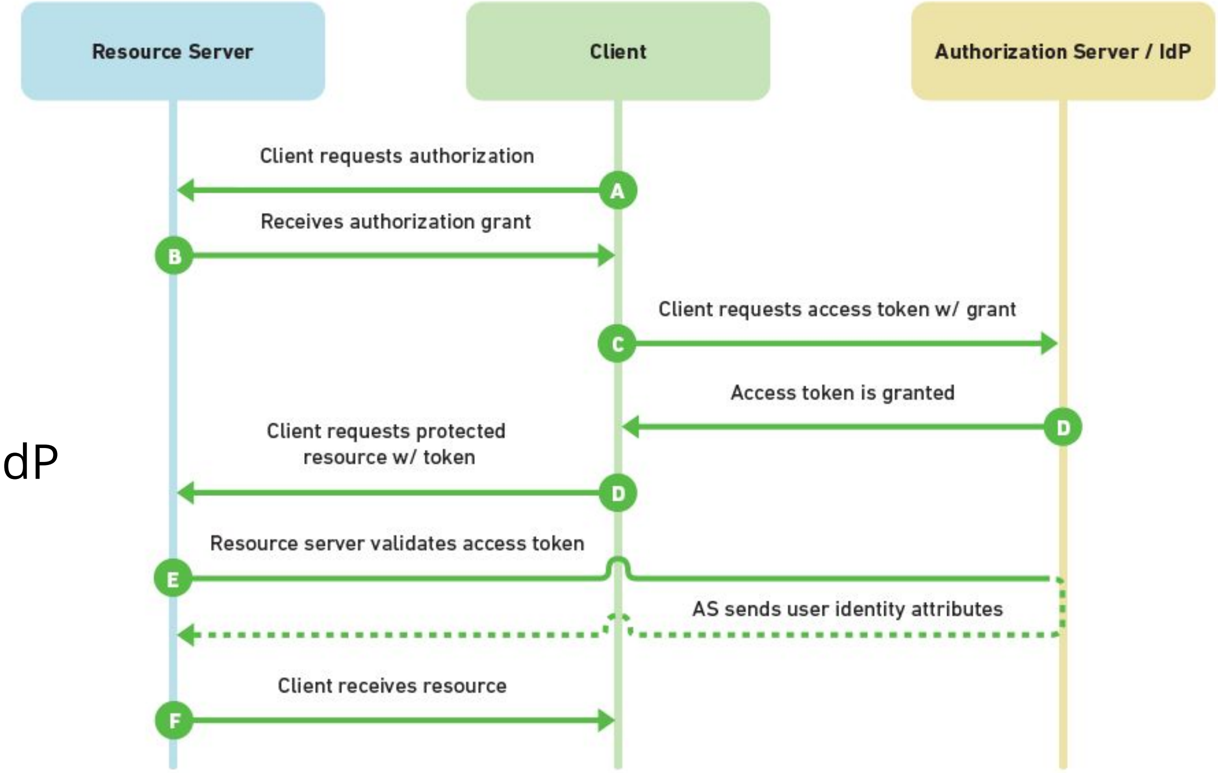


```
GET /oauthplayground/?code=4/sxr79FWJ_SPc-u5JcLMBnNIzAvN3.UnY1LhOM6UYZg
rKXntQAax2GVNw5fAI HTTP/1.1
Host: developers.google.com
```

https://code.tutsplus.com/articles/oauth-20-the-good-the-bad-the-ugly--net-33216

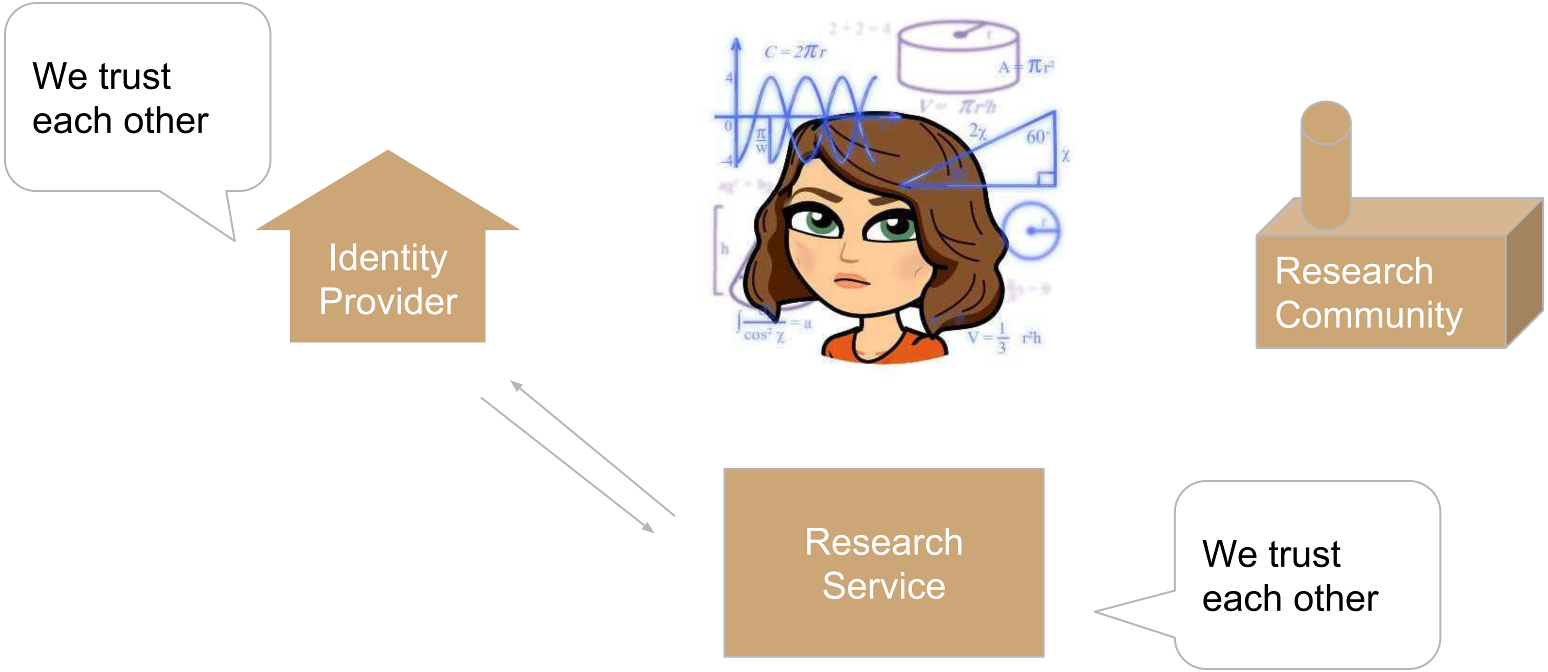
OAuth2 Protocol

- Client
  - User on their browser
- Resource Server
  - Website requiring Authorization
- Authorization Server/IdP
  - Home Organization


Resource Server
Client
Authorization Server / IdP
Client requests authorization
A
Receives authorization grant
B
Client requests access token w/ grant
C
Access token is granted
D
Client requests protected resource w/ token
D
Resource server validates access token
E
AS sends user identity attributes
Client receives resource
F

https://www.mutuallyhuman.com/blog/2013/05/09/choosing-an-sso-strategy-saml-vs-oauth2/

Where's the trust?



We trust each other

Identity Provider

Research Community

Research Service

We trust each other

OAuth2

Good Bits

- Tokens widely accepted
- Easy to implement
- Works for non-web

Bad Bits

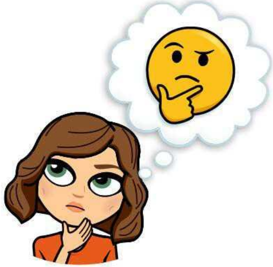
- Current identity federation status immature

A Quick Pit-Stop

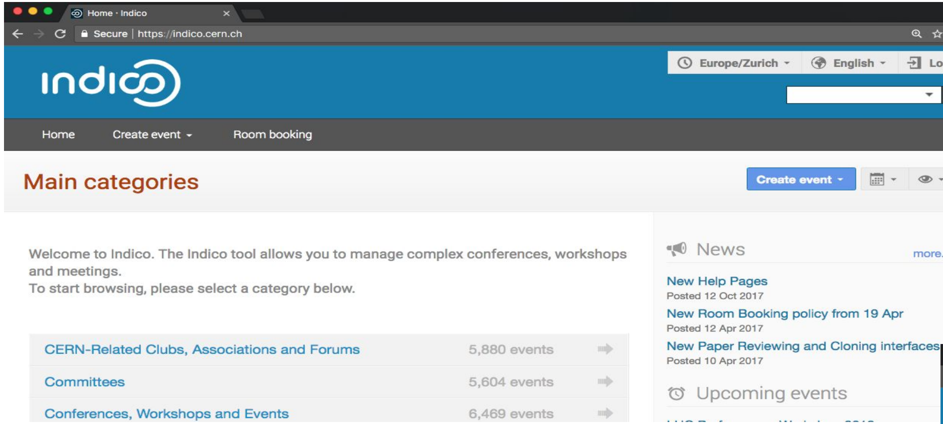
We looked at 3 different technologies for distributed authentication.

	Certificates	**SAML**	**OAuth2**
Web?	Yes	Yes	Yes
Command Line?	Yes	No	Yes
Advantage?	Simple	Scalability	Widely accepted
Disadvantage?	Security & Usability	Usability, Non-web	Scalability
Example	Grid Certificates	Your Home Organisation	ORCID, Github

Which one is best?



Brief Aside: Example at CERN



Home · Indico

Secure | https://indico.cern.ch

Europe/Zurich ▾ English ▾ Login

Home Create event ▾ Room booking

Main categories

Create event ▾

Welcome to Indico. The Indico tool allows you to manage complex conferences, workshops and meetings.
To start browsing, please select a category below.

CERN-Related Clubs, Associations and Forums	5,880 events	➡
Committees	5,604 events	➡
Conferences, Workshops and Events	6,469 events	➡

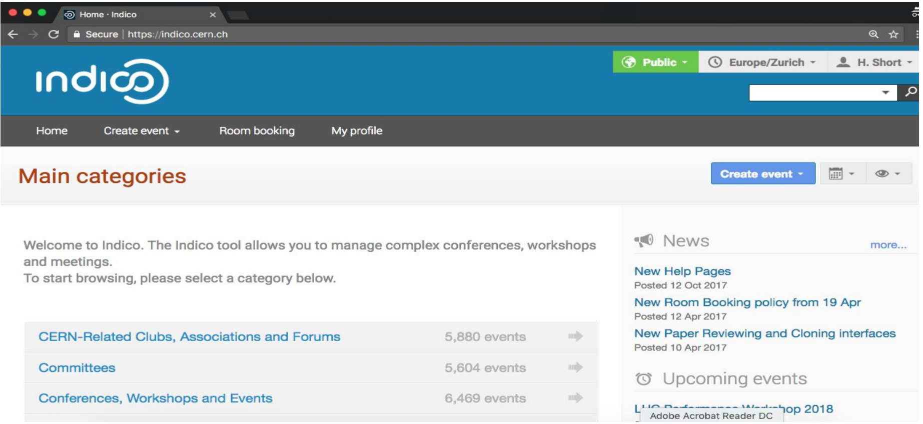
News more...

New Help Pages
Posted 12 Oct 2017

New Room Booking policy from 19 Apr
Posted 12 Apr 2017

New Paper Reviewing and Cloning interfaces
Posted 10 Apr 2017

Upcoming events



Home · Indico

Secure | https://indico.cern.ch

Public ▾ Europe/Zurich ▾ H. Short ▾

Home Create event ▾ Room booking My profile

Main categories

Create event ▾

Welcome to Indico. The Indico tool allows you to manage complex conferences, workshops and meetings.
To start browsing, please select a category below.

CERN-Related Clubs, Associations and Forums	5,880 events	➡
Committees	5,604 events	➡
Conferences, Workshops and Events	6,469 events	➡

News more...

New Help Pages
Posted 12 Oct 2017

New Room Booking policy from 19 Apr
Posted 12 Apr 2017

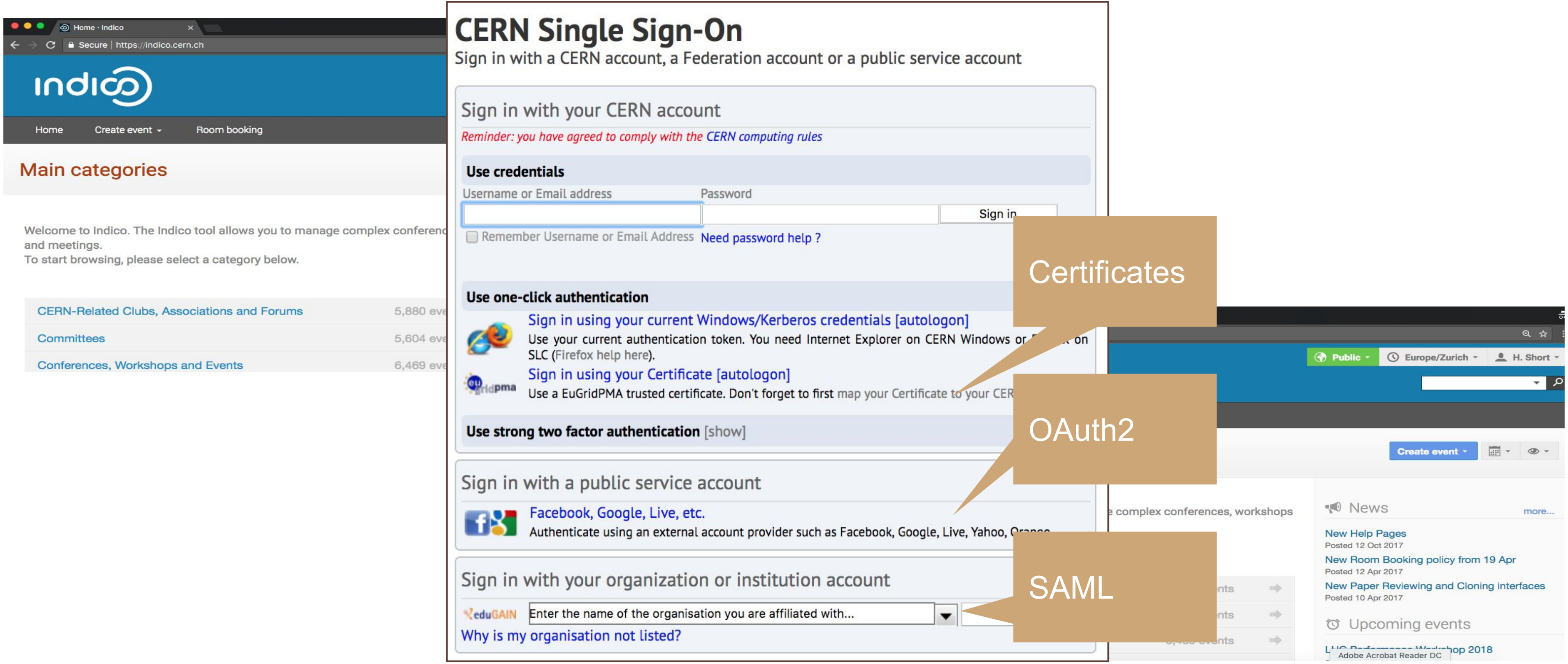
New Paper Reviewing and Cloning interfaces
Posted 10 Apr 2017

Upcoming events

LHC Performance Workshop 2018

Adobe Acrobat Reader DC

Brief Aside: Example at CERN



Home · Indico

Secure | https://indico.cern.ch

indico

Home   Create event ▾   Room booking

Main categories

Welcome to Indico. The Indico tool allows you to manage complex conferenc
and meetings.
To start browsing, please select a category below.

CERN-Related Clubs, Associations and Forums	5,880 eve
Committees	5,604 eve
Conferences, Workshops and Events	6,469 eve

CERN Single Sign-On

Sign in with a CERN account, a Federation account or a public service account

Sign in with your CERN account

Reminder: you have agreed to comply with the CERN computing rules

Use credentials

Username or Email address   Password

Sign in

□ Remember Username or Email Address Need password help ?

Use one-click authentication

Sign in using your current Windows/Kerberos credentials [autologon]

Use your current authentication token. You need Internet Explorer on CERN Windows or F   on SLC (Firefox help here).

Sign in using your Certificate [autologon]

Use a EuGridPMA trusted certificate. Don't forget to first map your Certificate to your CER

Use strong two factor authentication [show]

Sign in with a public service account

Facebook, Google, Live, etc.

Authenticate using an external account provider such as Facebook, Google, Live, Yahoo, Orange

Sign in with your organization or institution account

eduGAIN Enter the name of the organisation you are affiliated with... ▾

Why is my organisation not listed?

Public ▾   Europe/Zurich ▾   H. Short ▾

Create event ▾

e complex conferences, workshops

News   more...

New Help Pages
Posted 12 Oct 2017

New Room Booking policy from 19 Apr
Posted 12 Apr 2017

New Paper Reviewing and Cloning interfaces
Posted 10 Apr 2017

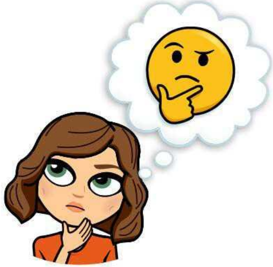
Upcoming events

LHC Performance Workshop 2018

Adobe Acrobat Reader DC

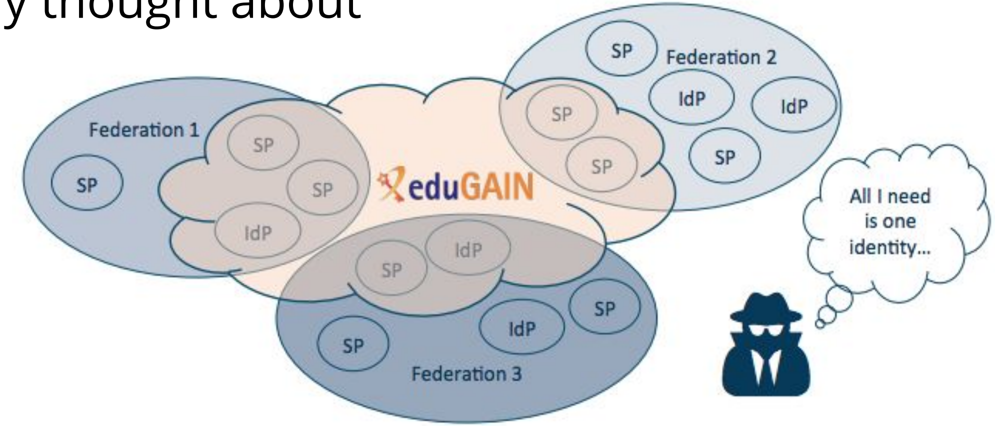
Security Incident Response

If an identity is compromised, how can we protect the community?



Security Incident Response

- What is Security Incident Response?
- Large communities
  - In eduGAIN > 4000 organisations
- Don’t necessarily know or trust each other
- Even if protocol is secure, nobody thought about security incident response :(



Suspension

- Each **service** could suspend the account
  - How can they share the information between each other?
  - How do they know when the compromise has been resolved?
  - Won’t this take a long time anyway?
- The **identity provider** could suspend the account
  - What if they don’t react quickly?
  - How do we contact them?
  - What if they refuse?

Investigation

- Will evidence be kept?
- Can logs be shared legally?
- Are contact points provided?
- If I share data, will the recipient respect confidentiality?

...

If nobody has thought about it, generally the answer is “No”

What can we do?

WLCG Certificate Federation

- Common security policies
- Central suspension mechanism (Argus)
- Infrastructure CSIRT (Computer Security Incident Response Team)

Very mature setup with international participation in trust initiatives (IGTF)

SAML Federations

- Newly established Security Framework
- No central suspension mechanism
- No central operational security or incident response capability

Still a long way to go before Research Communities trust them to the same extent

What am I hoping that you will remember?

- The global research community is increasingly connected through shared use of digital identities
- This brings benefits and also challenges
- There are multiple ways of doing it

If you are developing a service that needs authentication or authorisation, come back and look at this!

Questions?

Thanks for listening :)