# Notes FIM4R 21st Workshop Denver 2025

## Welcome and round of introductions

*Please check our spelling and correct if it's wrong! We did our best! Also feel free to embelish with your creativity :-)*

1. Maarten Kremers - SURF
2. Anders Sjöström- from Sweden (LU, SUNET, NORDUnet) - fighter for federated assurance
3. David Groep - from Nikhef and Maastricht University, Netherlands - e-Infrastructures, AARC & subject to policy
4. Dave Kelsey - STFC RAL UK - festive wise man of FIM
5. Hannah Short - CERN (my creativity has abandoned me now...)
6. Chris Wahlen - NIH
7. Tom Dack - STFC RAL
8. Liam Atherton - STFC RAL
9. Eisaku Sakane - Japan national institute of informatics HPCI gakunin federation
10. Waldo Fouche - AAF
11. Björn Mattsson - SUNET/SWAMID
12. Pål Axelsson - SUNET/SWAMID
13. Romy Bolton - Internet2
14. Mihály Héder - Hungary / GÉANT Project
15. Jody Tracy - Internet2
16. James Cramton - internet2
17. Russell Ianniello - AAF
18. Jim Basney - CiLogin
19. Mark Rank - Cirrus Identity
20. Scotty Strachan - Nevada System of Higher Education
21. Dana Brunson - Internet2
22. Matthew Puku - AAF (not registered?)
23. Chris Phillips - VeriMe.coop
24. Alan Buxey - Myunidays
25. Dalia Abraham - AAF
26. Albert Wu - InCommon/Internet2
27. Cory Snavely - NERSC, LBNL - several well funded new initiatives so seeking input
28. Kyle Lewis - NIH (mostly) & assurance evangelist
29. Ben Oshrin - Spherical Cow - likes trains
30. Ioannis Igoumenos - developer

31. Matthew Exonomou (not registered) - developer - fighting the good fight to mark it all easier!
32. Phil Smart - Jisc, UKAMF
33. Rob Smith - OpenAthens (not registered)
34. Shannon Rody - lawrence berkley lab
35. John Willis - Texas A&M University - looking to better integrate in incommon as only just starting
36. Michael Foggart (not registered)
37. Ann West - Internet2
38. Mike Jones - Microsoft

+5 who arrived after introductions!

No problem if you're not registered - it's just to make sure we record who came ;-)

# Internet2: supporting US research collaborations

- The team engages with Research Computing and Data people at communities, much experience
- Researchers want to focus on their research, and the landscape is a minefield
- Research Computing & Data (RCD) Professionals - a new profession
    - Understand the technology & the science, and can communicate in both spaces
- Faculty are expected to know many skills, e.g. AI, and the learning curve is large
- This matches very nicely to what we see in the AARC project, research communities need consultancy to get started
- CaRCC recently surpassed 2000 network members
- Collaborate to bridge the RCD & FIM4R communities - find the low hanging fruit and commonalities

## Nevada Case Study - Scotty Strachan, Research-IT Strategy Evangelist

Out the era where a single PI can make things happen - multiple organisations collaborating on highly technical topics.

No scientist can avoid technology these days.
"Research IT" individuals are distributed and embedded in "normal" campus IT teams
Several decades of technical debt incurred by US Higher Institution IT - as HE becomes a business, the IT becomes a business tool (mostly an IT support setup).
KPI is student enrollment rather than quality or overall IT architecture benefit
Network needs edge support for field science activities.
Unlikely that the person told to create the tech stack (often a scientist) has the necesary skills to do it well or securely.
Focus on enterprise "lights on" leads to falling behind in other ares - such as Federation-last

thinking.
Small Admin teams to support IT activities, with high turnovers
Needs:

- Clear & Well documented tooling
- Simplified Deployment Models
- Self-Service Hierarchy designs
- Centralised resource portal service
- Reliance on federated IDP for Auth (Federated First)

IAM systems are standing in the way of research and industry collaboration.
Onboarding services is painful due to lack of people
Idea of Research IAM as-a-service complete with connected applications (computer, data staging, sensors, vpn). CiLogon as highest level AAI, with a subsidiary Keycloak and grouper for delegated group management.
There's a lot of elements not in the diagram that are essential and might change.

Thought - seems to be a need for a service catalogue landing page for research AAIs

## Discussion

- Anders
  - Career path for TA staff - how do you find, and keep people with a complete career path.
  - Tools and resources for Career paths have been looked at and ideas beingmade available at the
  - manager of the Puhuri project. very complex integration model with multiple campuses
- Dave
  - somebody needs to provide tools to make it easy - should this be Internet2, or should it be somewhere else?
    - Scotty: SME expertise is spread across, but typically not engineers. People are engaged with the research process, and not the IT stack. What is missing is a solid concept of "what is Research IT, & R-IT Engineering". Not about uptime, but how fast can the science be done, and how flexible can you be.
  - Should this be international?
    - Dana: a lot of this is solved the the large International groups, as they have the IT resources. How do we use this experience to fix and support the mid to small groups and projects
- Hannah
  - Support availability? Reliance on AI Systems?
    - Biggest question with RCD community at large.

        - Support at the Dept level will always be in question, as it's never been a priority
        - National Compute Services are "not coming to save us" - build solutions with the minimal resources and support we can put together. Building infrastructures that can be maintained through staffing changes
    - In Europe there is often a "host university", that others connect into - is the same concept here?
        - Here, it's who "did it first"
        - In domain science, the PI often doesn't have the resources correct for data - let alone the rest of the IT stack
        - Sells the science short, things become Vaporware and processes just have to be repeared next time.
- John W
    - just beginning to allocate staff to support IAM infrastructure but at risk of losing them. If there were a professional track this may help with retention
        - Common job descriptions help, with support from HR
        - Certifications is next step for the RCD professions
        - Virtual residency trainings (and some others) that are generally accepted by the community
- Ann
    - Central IT will be doing the assurance processes and then asserting them
    - Researchers may or may not have other roles at universities
    - Trying to educate campus IT on how they can make researchers' lives better
        - Yes we need to continually evangelise with campus IT but that is not normally what they do and they use different terminology

# CILogon: supporting US research collaborations

Happy birthday CILogon - Celebrating 15+ years of effort to enable secure logon to scientific cyberinfrastructure (CI)

Consistent access control model across all connected services. Has grown over the years, 800+ organisations. Based on CoManage (this is missing from the AARC compendium).
Supports multiple token flavours for different communities: OIDC ID Tokens (eg ACCESS), SciTokens (eg LIGO), WLCG Tokens (eg Fermilab), GA4GH Passports (eg Australian BioCommons)
Aiming for sustainability, address technical debt. Plan for people moving on.
Saw a big increase when started OIDC support.
Now a fully subscription service (first tier is free).
X.509 removed quickly as large technical debt
Comange upgrade will be big

## Case study OSG

70 instututes participate, use the hosted comanage registry - support for nearly 200 research projects and over 400 individual members in the consortium. Have their own small team but need to distribute the admin work internationally.

They use ALL the pieces, including eduGAIN, other campus IdPs - users mustn't be blocked at the login stage. Use github flows. Link the ORCID in for data publication.
SSH keys uploaded into the registry. Also OAuth device flow.
They provide documentation with screenshots
OAuth Device Flow for SSH is good for security and user friendliness. Use the pam module from STFC.

## Discussion

- Hannah
  - how much of the infra setup is done by CILogin? E.g. the pam modules and SSH provisioning?
  - requests for full research community "branding"?
    - Yes, we put a lot of effort into the branding
  - SSH as a "backup" login during downtime for command line flows? CERN supports both SSH and Device Flow - has been very important for the team during disaster recovery
    - Can stack PAM modules - have a "break glass" PAM module for when the user ones don't work
- John - question about visual branding
  - Quite complicated as the CILogin logo is the one in the metadata
  - Branding exists in every piece of the puzzle - how can the community get the pieces into every tier
- Cory - best pratices and protocol support guidelines - how does CILogon adopt them?
  - All of CILogon's assurance support is consistent, passed along to downstream applications
- Mike - have the attributes been registered?
  - Yes, in IANA

# AARC: progress on AAI guidance for global research collaborations

AARC-Tree is the third iteration, and ends in February 2026.
Several (hopefully useful!) outputs over the next few months:

- Policy Development Kit Version 2
  - First version published a few years ago
  - losts of feedback from AAF, and have since aimed to simplify

    - Now 8 key steps to follow, with 6 key documents - ease of interoperability and support making policy declarations for the AAI
    - 8 Steps can be found on the wiki:
        - https://wiki.geant.org/spaces/AARC/pages/1214906481/AARC+PDK+v2
    - 6 policy documents
        - Attribute Authority Operational Security Policy
        - Acceptable Use Policy
        - Incident Response Procedure
        - Membership Management
        - Privacy Policy
        - Security Operational Baseline
- AARC Compendium of Best Practices & Recommendations
    - Builds on Research Infrastructure Use Cases based on stakeholder interviewers
    - Encourages community input on an ongoing basis, for Research Community Management, AAI Implementers & Operators, and Funding Agencies
    - Key Messages:
        - There is a benefit to common AAI solutions - bit to funding, and research collaborations
        - AAI complex, for many the recommendation is to use a hosted/managed solution. This aids future interoperability
        - A "consulting" service where Research Collaborations can seek advice
    - Please provide feedback at the doc!
        - https://docs.google.com/document/d/1buSq_L_rAW_C8ZZuTKQyjh7mldCduhOBRsXMMvPBRMg/edit?tab=t.0#heading=h.tj8rhc1rmy88

## AARC Outputs in Practice - European Open Science Cloud (EOSC)

AAI task force making use of AARC technical and policy guidelines to provide a gingle sign-on experience across multiple nodes (national or by research domain)
Nodes should trust each other - perfect match for OIDC Federation (Proxied token introspection as a temporary solution)

## Discussion

- Matthew: How many person weeks does it take to implement the median AARC blueprint, and how low do you think this would be in the future
    - Depends on the people: if people know what they are doing ~a few months, but if people are starting from scratch it could take years - becomes tempting to just use MS or Google
    - Scotty: The "time to science" should be a perfomance indicator. Frame the answers in this context - rather than forcing researchers to include this in their proposal, instead meet where they are at.

    - Chris P: why start from zero? Don't have to build it all your self. Acceptance/sales job question - ensuring information to guide decisions is crucial
    - CILogons free tier helps get the first buy-in
- Anders: not all doom and gloom, don't need an expert at every site. Can have communities attached to nodes, and build on that connection
    - EOSC onboarding is an online form, and is speedy

# AAF: progress in provision of federated identity services to Australian research collaborations

AAF has been around for ~16 years
25 NCRIS facilities, 90,000 researchers, 380+ international service connections and 12M authentications.
The plumbing works - but many researchers don't care until it breaks.
Core Theme: *Collaborations are still forming their identity strategy. Unsure where they want to go...*
Identity is an afterthought, inconvenient policy, and assumed trust. How to get the idea of a trust network to researchers. Researchers will solve their own problems, leading to siloing - AAF tries to uplift and assist here.
Why does this happen?

- Lack of governance,
- Don't understand Identity Assurance,
- Who's the community?,
- Collaboration Complexity

AAF PDK started from AARC and provides a template to get people going. Less mature collaborations struggle.

- high level governance (e.g. who can access)
- a list of approximately 9 docs

Collaborations in "incubators" that either have a reference architecture document created for them or use a hosted instance.
4 year plan to introduce an AAI platform to link different groups - a nation level solution

REMS (auto provisioning of entitlements) and CILogon (for groups and claims) are well used and liked

Challenges:

- Non education partners & assurance for their identities (e.g. from microsoft - also they want integration of a service with their microsoft AAI to be easy)
- How MFA is signalled across eduGAIN - chicken & egg based on IdP & SP support
- People want ID for life, and often end up with gmail as a workaround

- Low maturity in governance and lack of understanding in policy to implement in a policy-first manner
  - Reality is that policy is second and they don't want to even read it - the technology is used to understand the policy
- local accounts required in downstream services

## Discussion

- Anders:The ones who need the assurance are the resource providers, who have the facilities. Do they show an interest in getting this info? From his experience, no interest was shown and the facilities didn't consider it - this has now completely changed, and assurance is much more important now
- Hannah: What kind of problems have you come across in implementing a Microsoft IdP within the platform?
  - Series of EntraID domains to be linked to the platform. Have given up on working with Microsofts SAML implementation, were running a Shibboleth proxy but have given up.
  - Shibboleth has been replaced by Satosa - gives a lot of control on "taking Microsoft's garbage", and then translating and tranforming this for the rest of the federation.
  - The answer to bringing Federated IdPs into Microsoft is hiding the federation between an IdP proxy
  - Request for an ACAMP session on Microsoft and Proxies
- Scotty & David: site local proxies (i.e. a mini proxy in front of a smaller set of services) is a common pattern
- Anders: we need to solve the problem of affiliation tokens so that it cannot be bypassed through proxies
  - Who you are, where are you from, and how sure are you about that?

## Open Space

- What do we want to bring to ACAMP?
  - Interacting with Microsoft Entra as a Research AAI (Hannah)
  - What features of OpenID Federation are we really excited about? (Phil) - very good time as review is open
  - What is OpenID Federation? (Mike)
  - Professional certification? (John?)
  - Bridging the gap between campus IT and research computing people - connect the people (Not Dana as she's gone) -> do workshop at PEARC (https://pearc.acm.org/pearc26/), in Minneapolis next July, could also do a virtual group in CaRCC (https://carcc.org/)
  - Shall we do a FIM4R v 3 paper? (Maarten)

- how can we better engage with research collaborations?
- engaging the campus IT as well as research PIs?
- John: received some government push to address IAM so it does seem to be taken seriously somehow
- Jim: the ID Pro organisation is becomming very mature
- Mihály: STM publishers created a recommendation for trusted verification (this is a group of highly trusted publishers)
    - 
    - In the incubator looking for new topics that aren't well covered
    - As well as assurance you need academic track record (bona fide)
        - **Here is the demo for the bona fide topic**: https://wiki.geant.org/spaces/G52W5/pages/1139736659/Cycle+10+final+Demo (first presentation)
    - Establish identity and then link various things to it
    - People were reviewing their own work by creating different identities
    - Other side issue where doctors were reviewing papers outside their domain
    - Many different publishers working under STM
        - Here is the STM recommendation titled **Trusted Identity in Academic Publishing | Part 2: The Researcher Identity Verification Framework**: https://stm-assoc.org/document/trusted-identity-in-academic-publishing-2/
    - The login flow for publishing is not painful these days, but the academic track record is complex
    - Matrix of assurance for "evidence of academic work"
    - This is open for comment at the moment
    - If large publishers are going to insist on this then our IdPs are going to have to start issuing new things
    - Cross-Ref (DOI) is relied on quite heavily, and ORCID
    - ORCID is becoming more important than eduGAIN IdPs - potentially a single global IdP. This may mean that we have to onboard all our researchers to ORCID for them to do their research.
    - The publishers themselves do manual ID checking
    - Participation very welcome!!!!
- Matthew: has anyone tried analysing the conversion of funding budget to scientific output?
    - Cory - with ORCID it is a bit easier
    - David - as a funding agency he asks for a list of outputs. Very difficult for big infrastructures (like WLCG) where there is no clear mapping
    - Anders - papers have to include funding string in the footer so you can at least search for it
    - Mihály - recommendation for acknowledgements is to use a specific schema to do it programatically and link to funding IDs. This is also mentioned in the STM paper.

There will also be author types introduced (e.g. main author, translator etc)

- Crossref Funder Registry: https://www.crossref.org/services/funder-registry/
- Contributer Role Taxonomy: https://credit.niso.org/

# Wrap up

# General vibes picked up during the day

- Research infrastructures are SO complex that we HAVE to make it easier. That means providing hosted (easy to use - technically and financially) research AAIs (potentially with a default set of connected resources)
- We should do a v3 on complex virtual research environment AAI requirements <- need to highlight the rationale for that complexity while recognizing reality of where HE IT is

# Notes for content for FIM4R 22

- Implementation of EOSC, what is difficult in reality? Keep it end user facing
  - Thematic node: CERN (H to ask tibor.simko@cern.ch?)
  - National node: SURF (Maarten)
  - Someone from EOSC association
  - Euro HPC
  - How does OIDFed look in reality (how will it look?)
  - Swedish one: Anders
- Use case of a smaller Research Collaboration (who?)
- OIDFed mapping to a real research use case (trust between AAIs)
- When do we stop trying to solve SAML problems? How do we map from one to the next?
- What to put in FIM4R v 3? (interactive sessions)
  - Progress from the other 2
- What if ORCID is the best solution?
- Account linking, is this a good solution?
- Re-assess the previous FIM4R papers
  - We did do a quiz... H to find the responses
  - What should we be prompting eduGAIN to do?
- Pains of attributes changing
- Present AARC things

Morning: Talks
Afternoon: FIM4R v 3 practical session