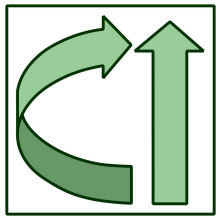
Propelling Australia's Digital Life Science Research with Collaborative Research Infrastructure



Jim Basney
Principal Research Scientist - CILogon

John Scullen
Head, Projects & Managed Services - Australian Access Federation

Background

Context: R&E in Australia

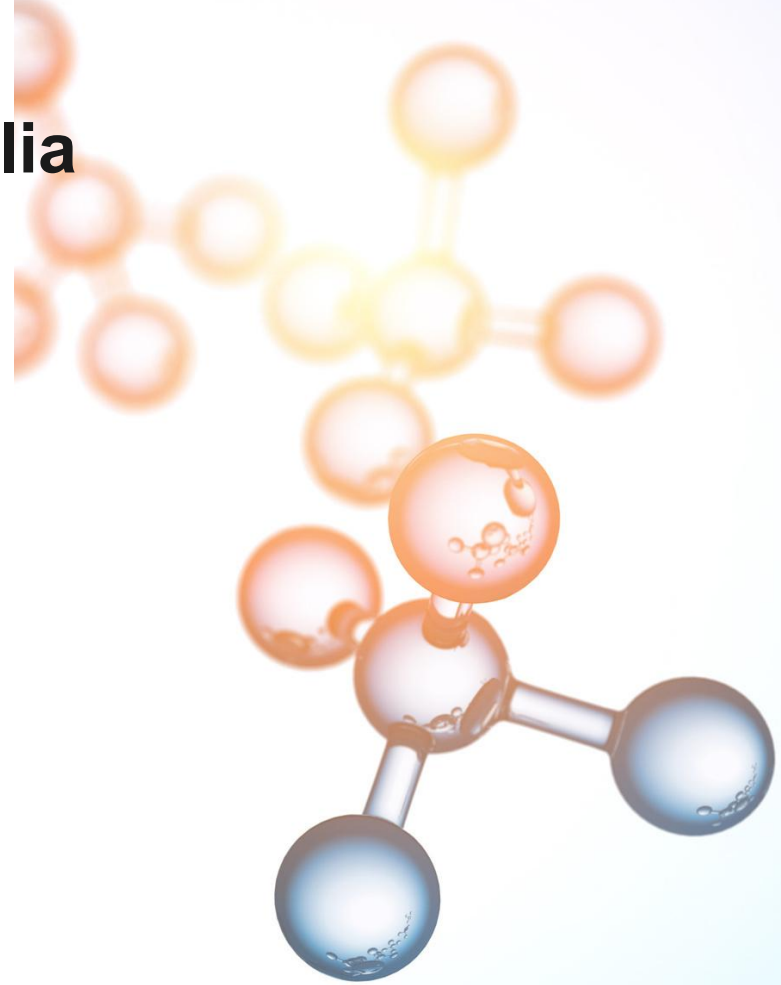
- **High penetration of enterprise solutions like Okta and Azure AD**
  - Largely addressed on-campus and cloud IDM challenges
  - Don’t handle federation (as we know it). >1 IdP = 🤯

- **Increasing need to include non-AAF research partners**
  - government agencies
  - hospitals
  - smaller medical research institutes
  - commercial providers

Context: R&E in Australia

Growing recognition of the need for improved security:

- more robust identity assurance
- authentication assurance (multi-factor authentication)
- signaling assurance levels between federation participants
- coordinated intelligence and incident response



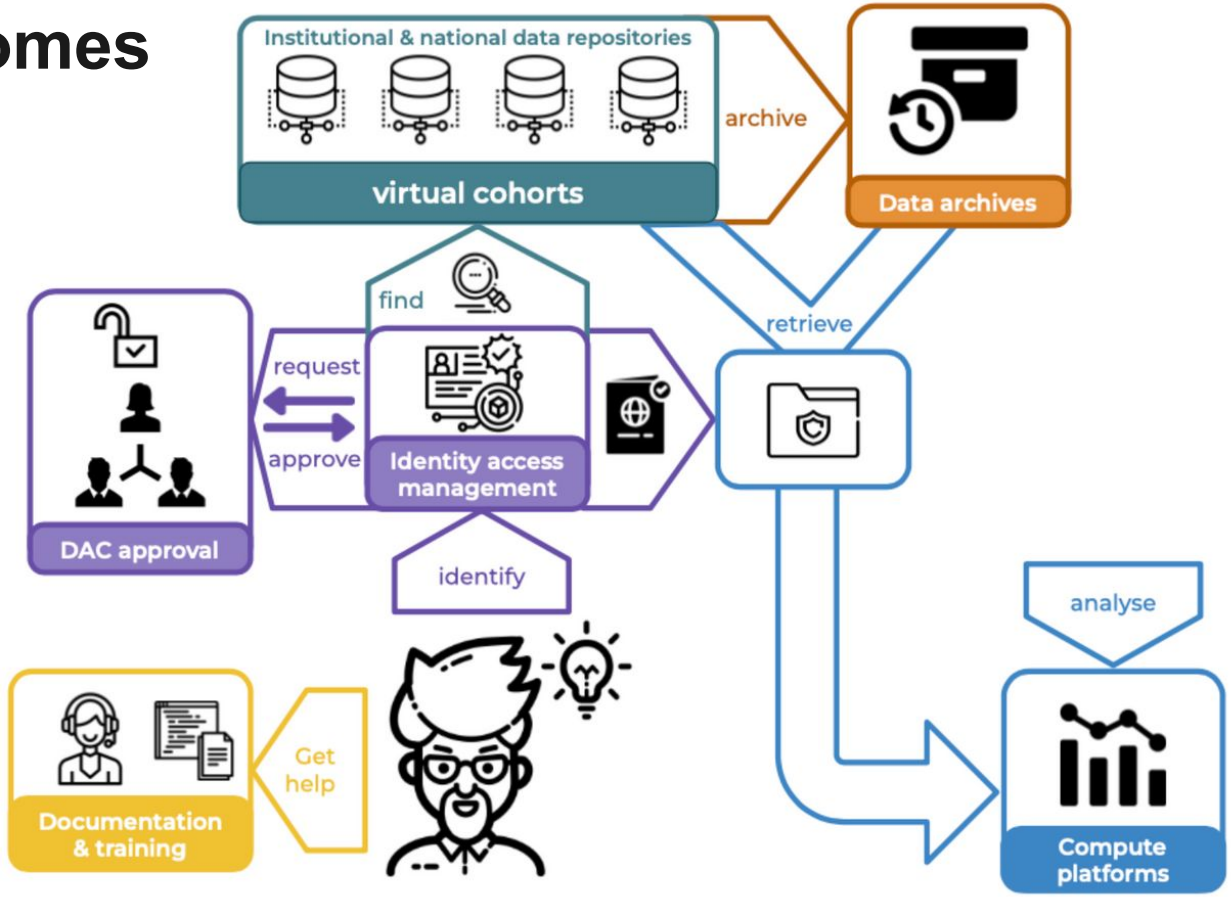
Australian BioCommons (think ELIXIR for Australia)

- **Mission: Enhancing Australia's digital life science research through world class collaborative distributed infrastructure**
- **30,000 life science researchers in Australia**
- **Approach**
  - Repurpose as much existing work as possible
  - Compatibility with international counterparts
  - Get the right experts together and work out what to do from there
  - Partner with existing providers to operate infrastructure

Pilot: Human Genomes Platform Project

Work Packages

1. Virtual cohorts
2. Streamline the Data Access Committee (DAC) process
3. Federated identity and access management
4. Data and metadata archiving
5. Documentation and training



HGPP Partners

Challenges

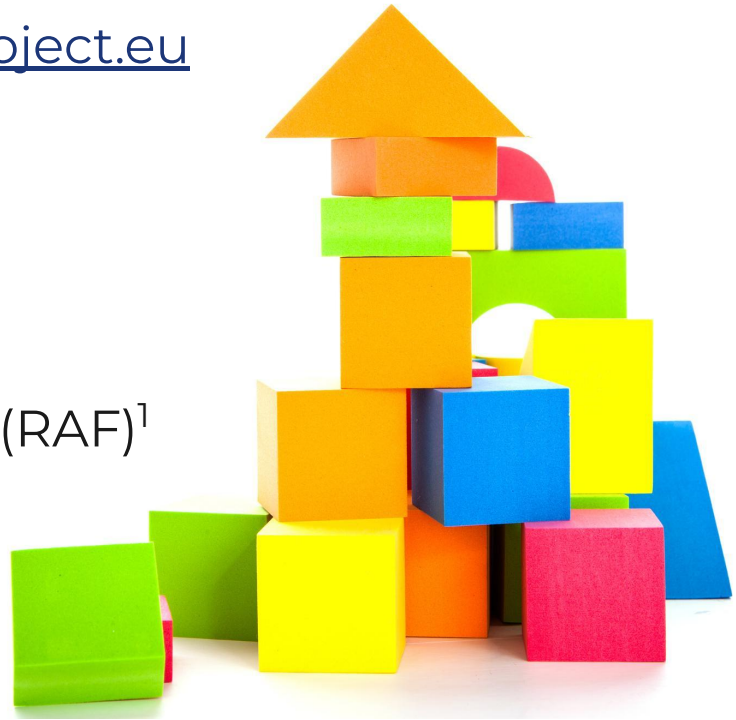
- **No F2F interaction**
- **Diverse collaboration networks**
  - most not AAF customers.
- **Diverse technologies**
  - OIDC, SAML
  - Web, command line, sensors and instruments
  - Need to build workflows between different service providers
- **No common rules or shared expectations re: trust and identity**

Challenges (continued)

- **Sensitive data (human genomics)**
  - catalyst for REFEDS Assurance Framework
- **Unfamiliar solution components – Gen3, Terra, REMS, CTRL, GA4GH Passports**
- **Difficult to coordinate Data Access Committees (DAC) approvals**
  - membership from multiple organisations
  - relies on email trails for approval
  - requests for further information add latency
  - process is opaque to the requestor and DAC members

Solution Building Blocks

- Start with AARC: https://aarc-project.eu
  - AARC Policy Development Kit
  - AARC Blueprint Architecture
    - do-it-yourself (DIY)
    - eduTEAMS
    - CILogon
- REFEDS Assurance Framework (RAF)[1]
- REFEDS MFA Profile

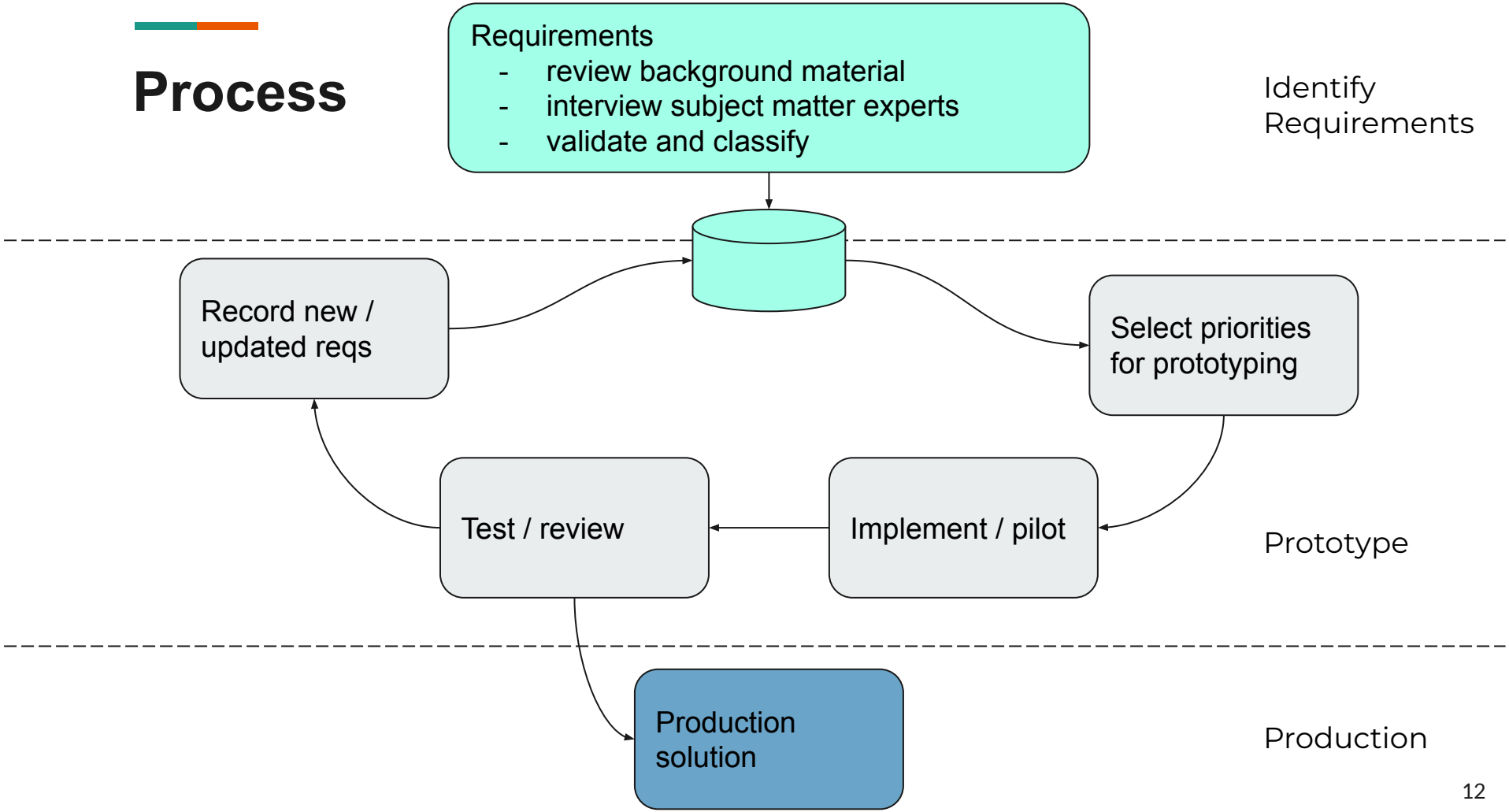


[1] https://refeds.org/assurance

Opportunities

- Growing recognition of AARC Blueprint Architecture for research communities and the role of tools like CILogon
- Growing awareness and acceptance of CILogon as a capable solution
- Reseller arrangement with CILogon
- AAF working to secure government funding to establish ongoing T&I capability

Process

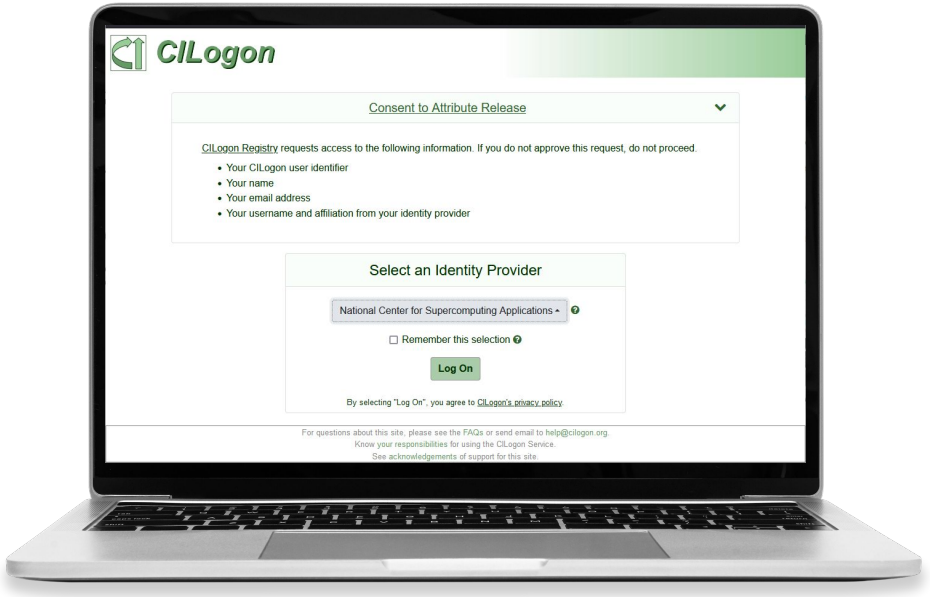

Requirements
- review background material
- interview subject matter experts
- validate and classify
Identify Requirements
Record new / updated reqs
Select priorities for prototyping
Test / review
Implement / pilot
Prototype
Production solution
Production

CILogon

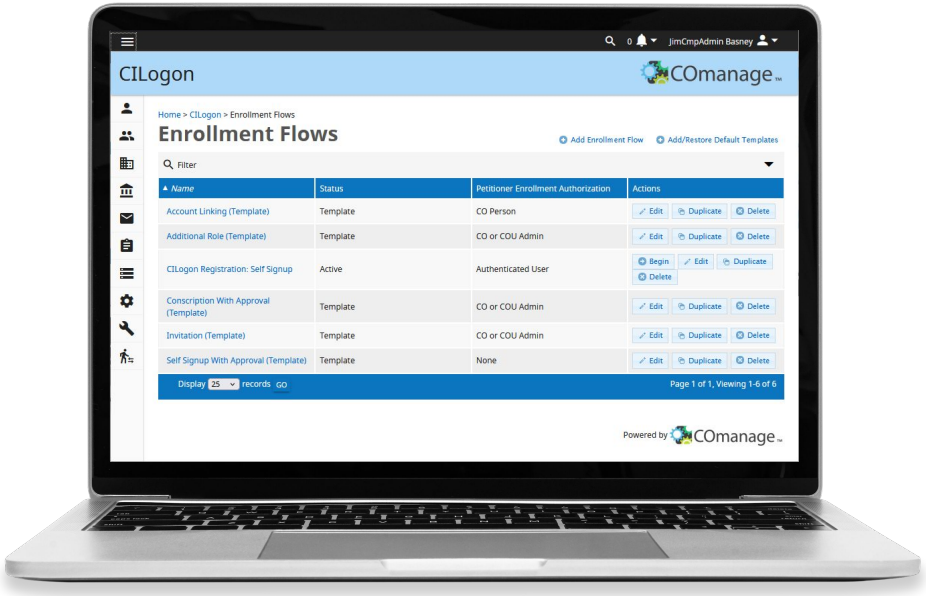
Introduction to CILogon

Federated identity management enables researchers to use their home organization identities to access research applications

- 17,500+ active users
- 450+ organizations globally



Introduction to CILogon (cont)



CILogon enables researchers to log on to cyberinfrastructure (CI) by providing coherent identity and access management for research collaborations

- via hosted cloud IAM services
- according to the AARC Blueprint Architecture

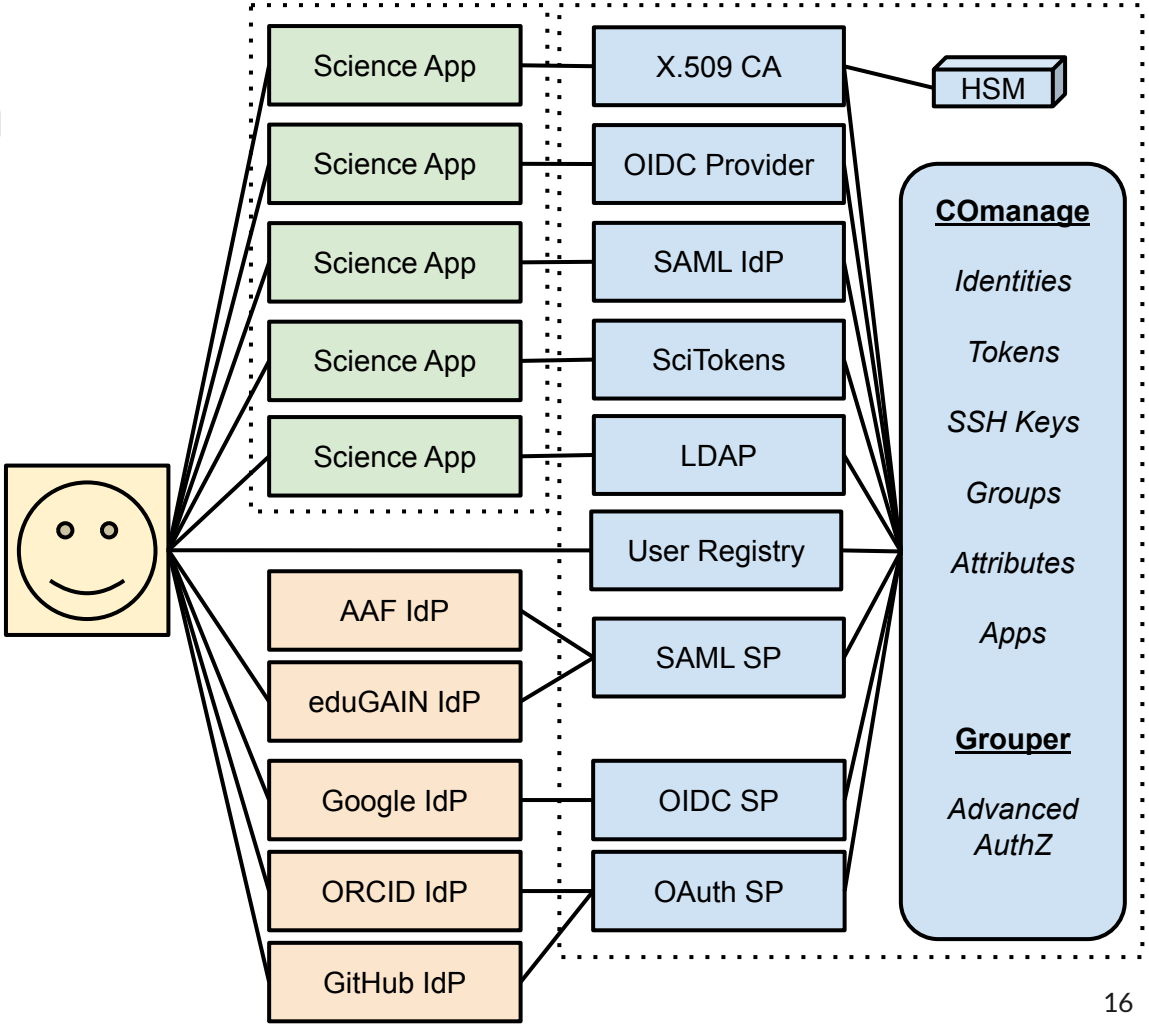
CILogon Platform

Supporting access to science applications:

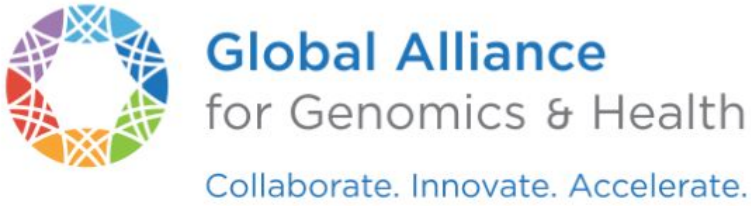
- HPC clusters
- Jupyter notebooks
- Globus
- REST APIs
- etc.

Identity providers from:

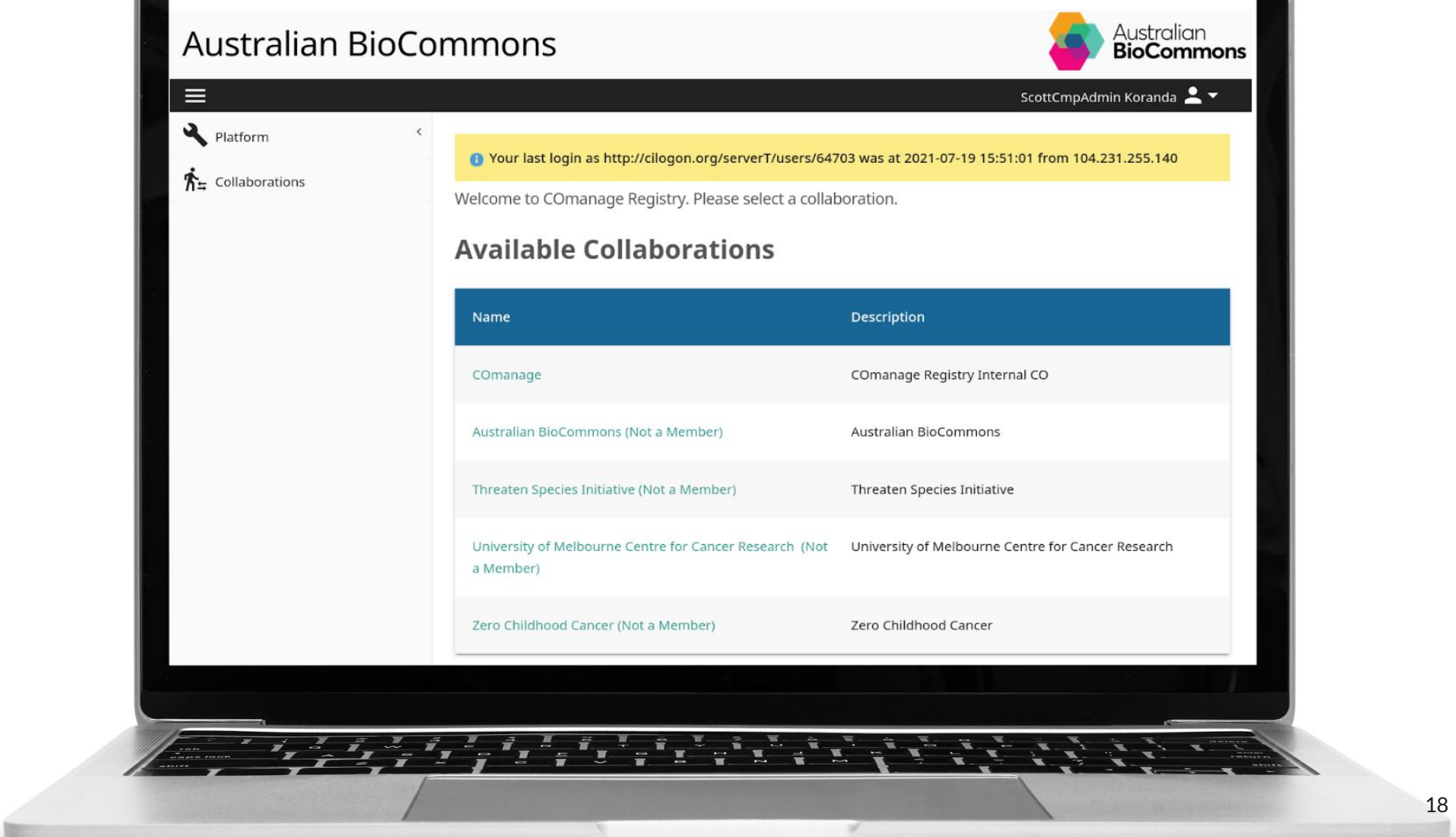
- Home organisations (SAML / Azure AD)
- external sources (Google, Microsoft, ORCID, Github)

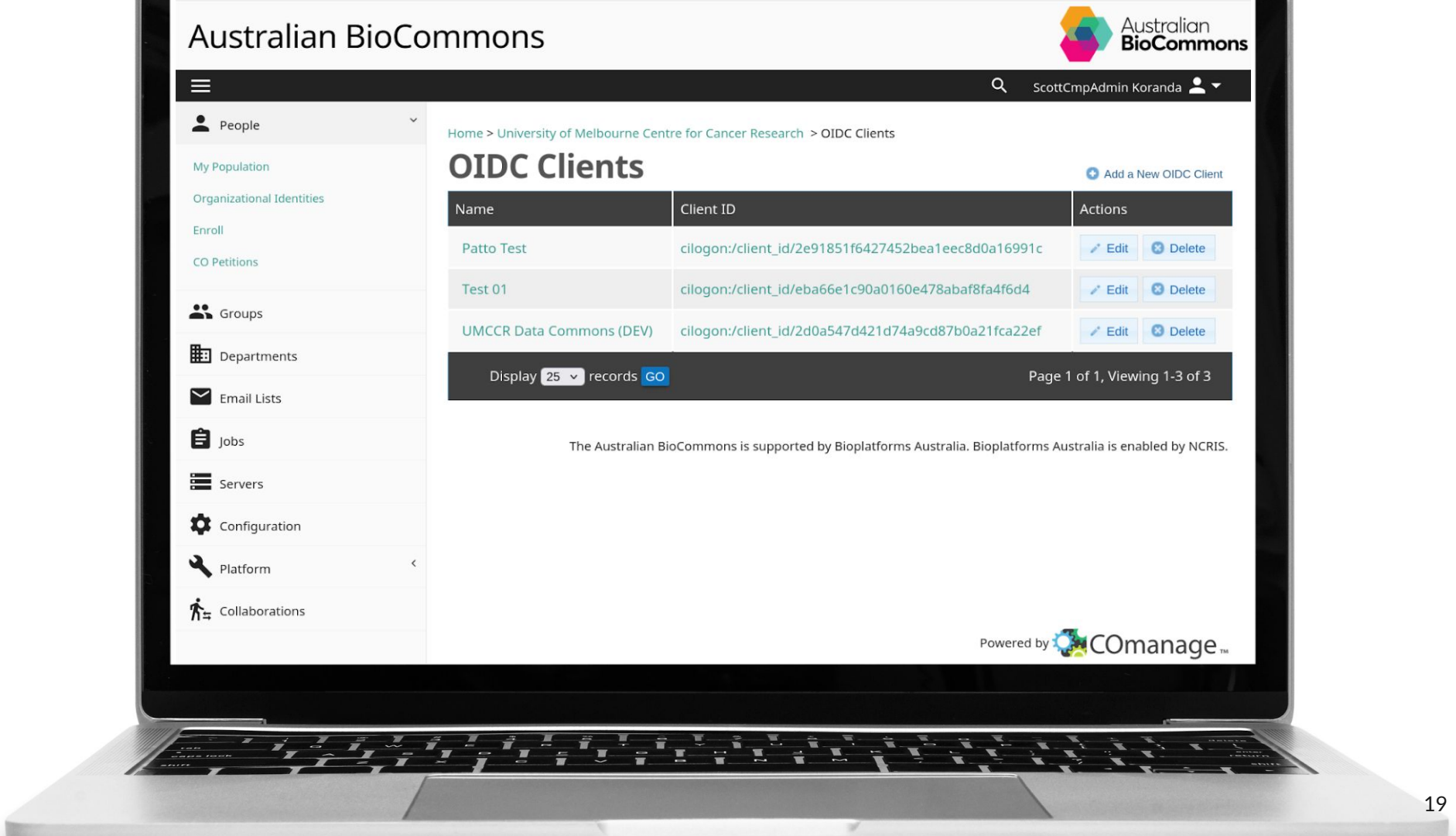


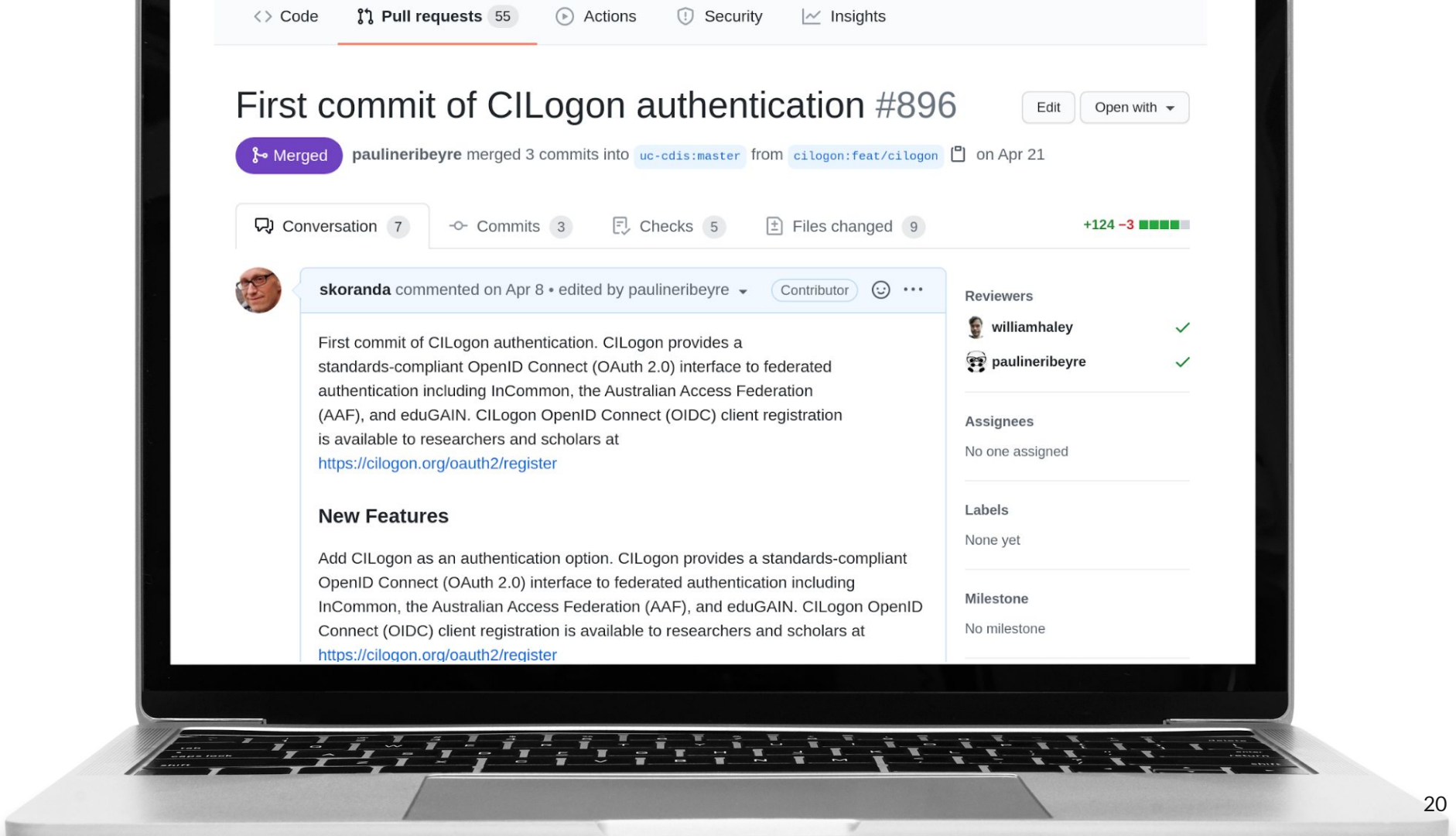
JSON Web Tokens for Science

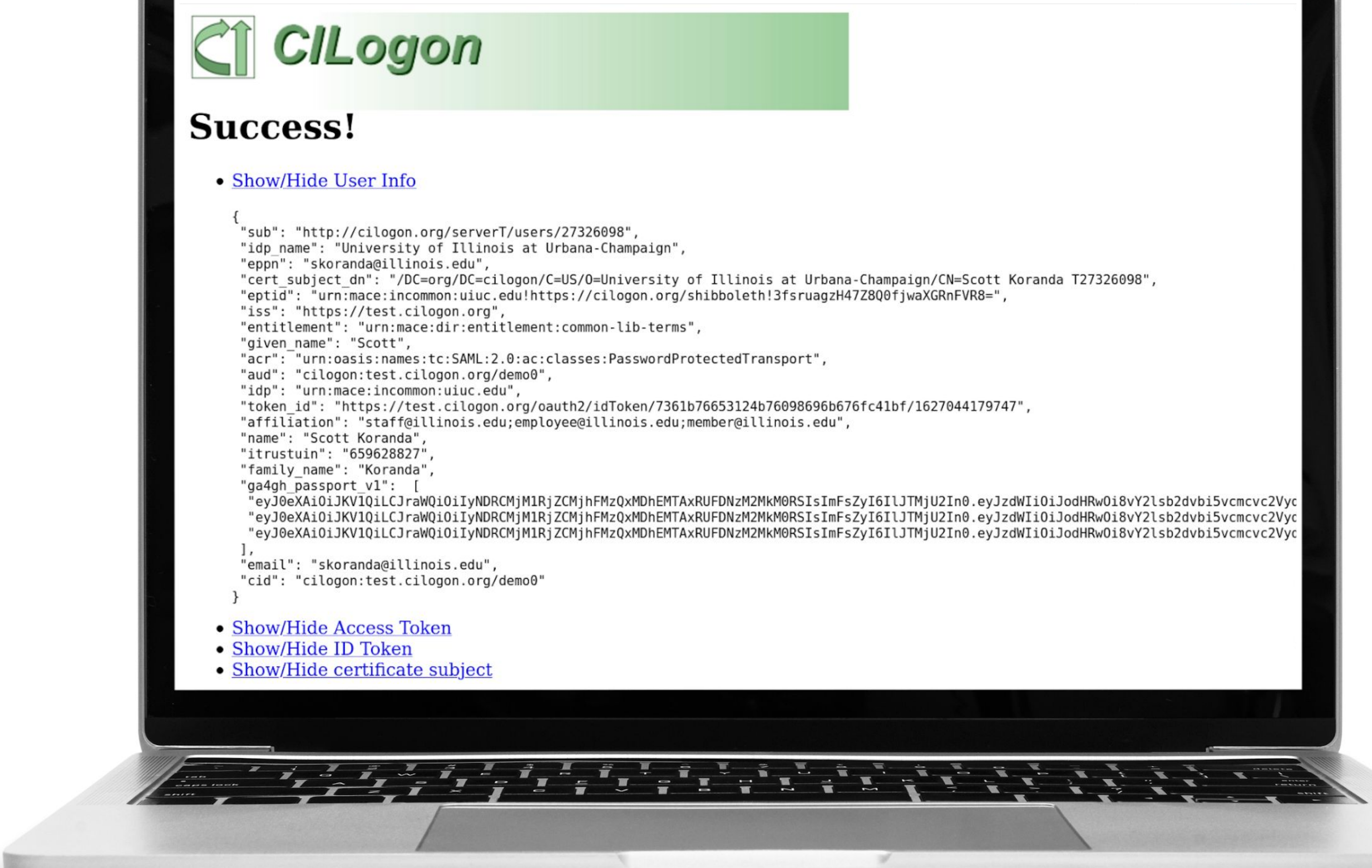


- containing user attributes and group memberships from the research community (via COmanage) and from the researcher's home institution via AAF / InCommon (eg. SCiMMA)
- containing authorization scope values determined by per client/subscriber policy (eg. LIGO)
- support for wlcg.groups and storage.*|compute.* scopes (eg. Fermilab)
- support for Affiliation And Role, Accepted Terms And Policies, ResearcherStatus, ControlledAccessGrants, and LinkedIdentities (eg. Australian BioCommons)

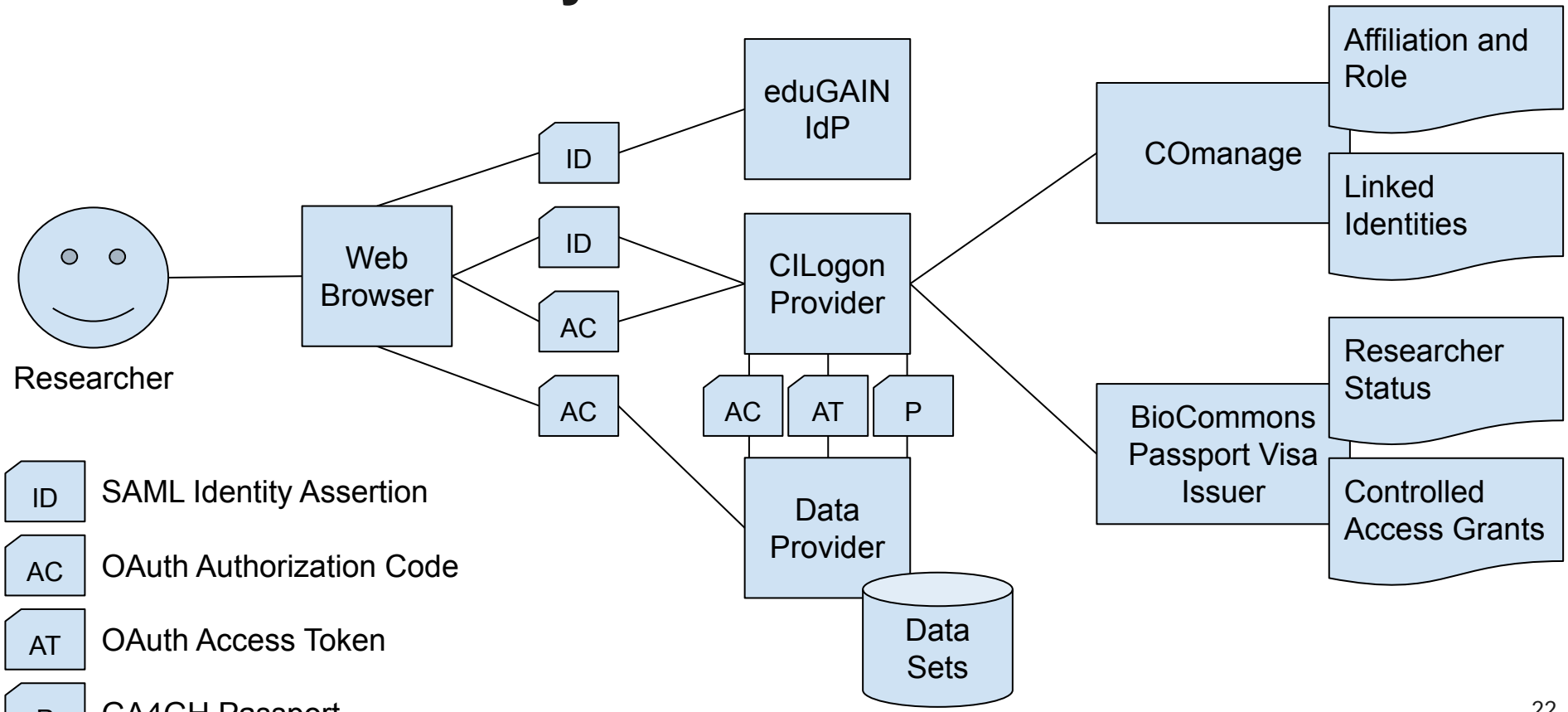

Australian BioCommons
Australian BioCommons
ScottCmpAdmin Koranda
Platform
Collaborations
Your last login as http://cilogon.org/serverT/users/64703 was at 2021-07-19 15:51:01 from 104.231.255.140
Welcome to COmanage Registry. Please select a collaboration.
Available Collaborations
Name
Description
COmanage
COmanage Registry Internal CO
Australian BioCommons (Not a Member)
Australian BioCommons
Threaten Species Initiative (Not a Member)
Threaten Species Initiative
University of Melbourne Centre for Cancer Research (Not a Member)
University of Melbourne Centre for Cancer Research
Zero Childhood Cancer (Not a Member)
Zero Childhood Cancer


Australian BioCommons
Australian BioCommons
ScottCmpAdmin Koranda
People
My Population
Organizational Identities
Enroll
CO Petitions
Groups
Departments
Email Lists
Jobs
Servers
Configuration
Platform
Collaborations
Home > University of Melbourne Centre for Cancer Research > OIDC Clients
OIDC Clients
Add a New OIDC Client
Name
Client ID
Actions
Patto Test
cilogon:/client_id/2e91851f6427452bea1eec8d0a16991c
Edit
Delete
Test 01
cilogon:/client_id/eba66e1c90a0160e478abaf8fa4f6d4
Edit
Delete
UMCCR Data Commons (DEV)
cilogon:/client_id/2d0a547d421d74a9cd87b0a21fca22ef
Edit
Delete
Display 25 records GO
Page 1 of 1, Viewing 1-3 of 3
The Australian BioCommons is supported by Bioplatforms Australia. Bioplatforms Australia is enabled by NCRIS.
Powered by COmanage


Code
Pull requests 55
Actions
Security
Insights
First commit of CILogon authentication #896
Edit
Open with
Merged
paulineribeyre merged 3 commits into uc-cdis:master from cilogon:feat/cilogon on Apr 21
Conversation 7
Commits 3
Checks 5
Files changed 9
+124 −3
skoranda commented on Apr 8 • edited by paulineribeyre
Contributor
First commit of CILogon authentication. CILogon provides a standards-compliant OpenID Connect (OAuth 2.0) interface to federated authentication including InCommon, the Australian Access Federation (AAF), and eduGAIN. CILogon OpenID Connect (OIDC) client registration is available to researchers and scholars at https://cilogon.org/oauth2/register
New Features
Add CILogon as an authentication option. CILogon provides a standards-compliant OpenID Connect (OAuth 2.0) interface to federated authentication including InCommon, the Australian Access Federation (AAF), and eduGAIN. CILogon OpenID Connect (OIDC) client registration is available to researchers and scholars at https://cilogon.org/oauth2/register
Reviewers
williamhaley
paulineribeyre
Assignees
No one assigned
Labels
None yet
Milestone
No milestone


CILogon
Success!
Show/Hide User Info
{
"sub": "http://cilogon.org/serverT/users/27326098",
"idp_name": "University of Illinois at Urbana-Champaign",
"eppn": "skoranda@illinois.edu",
"cert_subject_dn": "/DC=org/DC=cilogon/C=US/O=University of Illinois at Urbana-Champaign/CN=Scott Koranda T27326098",
"eptid": "urn:mace:incommon:uiuc.edu!https://cilogon.org/shibboleth!3fsruagzH47Z8Q0fjwaXGRnFVR8=",
"iss": "https://test.cilogon.org",
"entitlement": "urn:mace:dir:entitlement:common-lib-terms",
"given_name": "Scott",
"acr": "urn:oasis:names:tc:SAML:2.0:ac:classes:PasswordProtectedTransport",
"aud": "cilogon:test.cilogon.org/demo0",
"idp": "urn:mace:incommon:uiuc.edu",
"token_id": "https://test.cilogon.org/oauth2/idToken/7361b76653124b76098696b676fc41bf/1627044179747",
"affiliation": "staff@illinois.edu;employee@illinois.edu;member@illinois.edu",
"name": "Scott Koranda",
"itrustuin": "659628827",
"family_name": "Koranda",
"ga4gh_passport_v1": [
"eyJ0eXAiOiJKV1QiLCJraWQiOiIyNDRCMjM1RjZCMjhFMzQxMDhEMTAxRUFDNzM2MkM0RSIsImFsZyI6IlJTMjU2In0.eyJzdWIiOiJodHRwOi8vY2lsb2dvbi5vcmcvc2Vyc
"eyJ0eXAiOiJKV1QiLCJraWQiOiIyNDRCMjM1RjZCMjhFMzQxMDhEMTAxRUFDNzM2MkM0RSIsImFsZyI6IlJTMjU2In0.eyJzdWIiOiJodHRwOi8vY2lsb2dvbi5vcmcvc2Vyc
"eyJ0eXAiOiJKV1QiLCJraWQiOiIyNDRCMjM1RjZCMjhFMzQxMDhEMTAxRUFDNzM2MkM0RSIsImFsZyI6IlJTMjU2In0.eyJzdWIiOiJodHRwOi8vY2lsb2dvbi5vcmcvc2Vyc
],
"email": "skoranda@illinois.edu",
"cid": "cilogon:test.cilogon.org/demo0"
}
Show/Hide Access Token
Show/Hide ID Token
Show/Hide certificate subject

Demo: Thursday 12:45 - 13:15



Reseller Model

- AAF acts as local reseller and provides local support
- Infrastructure hosted in Australia

Roles

Research communities

- Manage community membership lifecycle
- Membership adds, moves and changes
- Connecting services to the collaboration

AAF

- Business analysis (process mapping for communities)
- Policy guidance
- Technical implementation and configuration
- Tier 1 and 2 support
- Advise CILogon about new feature priorities

CILogon

- Infrastructure maintenance
- New feature development
- Tier 3 support

Deployments

Life Science

BioCommons

- https://biocommons.org.au

Human Genomics

Cardiovascular Disease Research

Humanities

CADRE

- Sensitive social science data
- https://cadre5safes.org.au

Language Data Commons of Australia (LDaCA)

- Indigenous and other language collections
- https://www.ldaca.edu.au

Lessons

Lesson 1

Getting the first community established is the hardest part

- Abstract and hard to understand
- Demonstrators make it real
- Working examples make for a more compelling case

Lesson 2

Seeing how people interact with a basic prototype is more informative than endless navel-gazing

- Don’t waste time theorising about aspects that may turn out to be relatively unimportant
- The waterfall model doesn’t work well in research oriented projects.
- The research world has a much higher tolerance for uncertainty, ambiguity and experimentation than central IT environments. Leverage this to explore new technologies and techniques.

Lesson 3

Researchers need as much (more) help with the policy aspects as they do with technology

- Data protection & privacy, security incident response, and membership management lifecycle are often not even on the radar.
- Difficult to retrofit policy afterwards

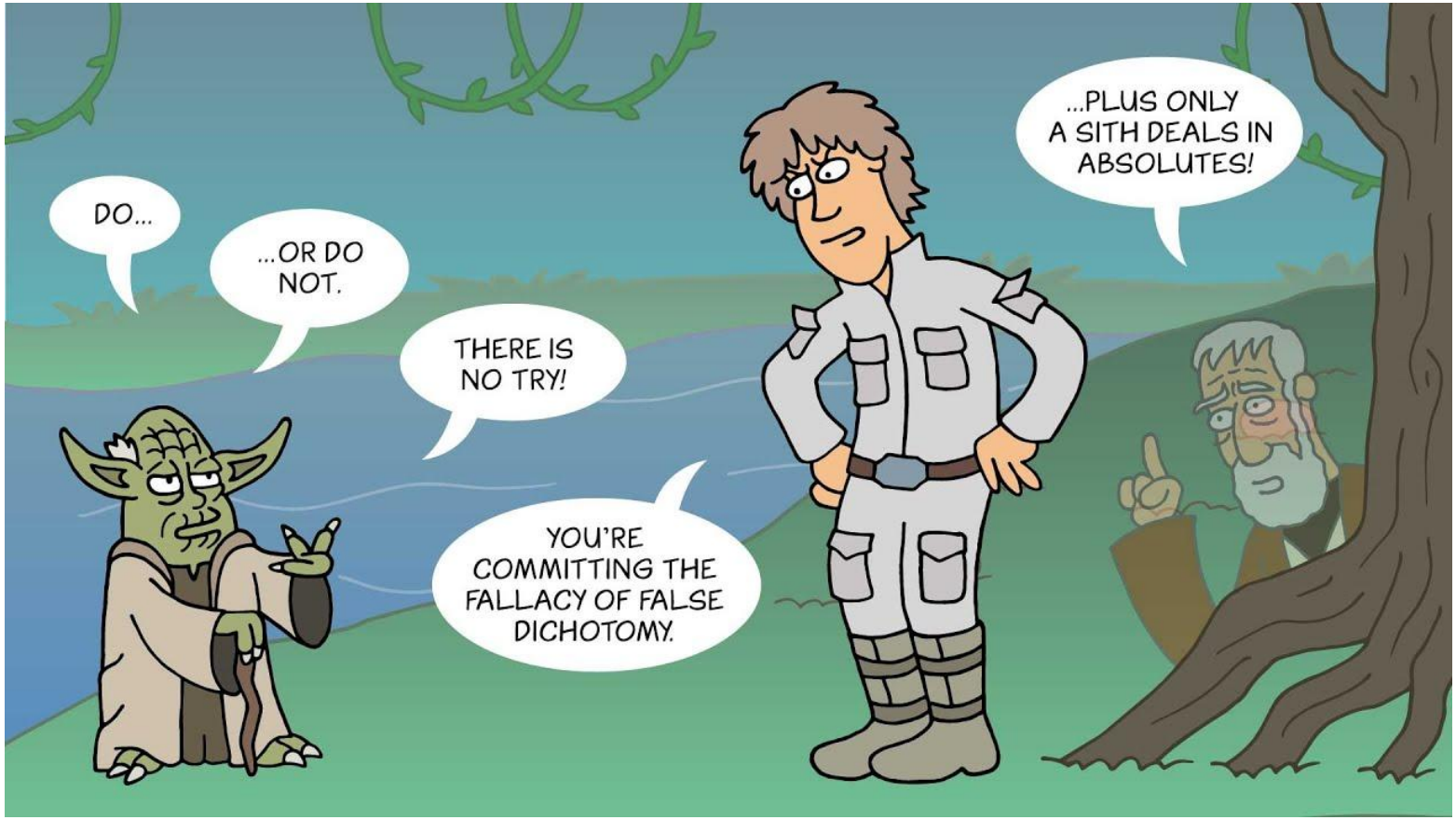
Lesson 4

Patterns transfer readily to other contexts / disciplines

- Cancer research to cardiovascular disease
- Transfer of ideas between life science and humanities communities

Lesson 5

Attributes and assertions vs access tokens... why not both?


DO...
...OR DO NOT.
THERE IS NO TRY!
YOU'RE COMMITTING THE FALLACY OF FALSE DICHOTOMY.
...PLUS ONLY A SITH DEALS IN ABSOLUTES!

Next Steps

From here…

- Confirm top priorities from each partner organisation
- Start building MVP
- Draft policies aligned with AARC Policy Development Kit and socialise
- Finalise reseller agreement details