

Enabling Federated Identity and Access Management for Scientific Collaborations

Jim Basney <jbasney@ncsa.illinois.edu> April 2022 MAGIC Meeting





IAM for Research Collaborations

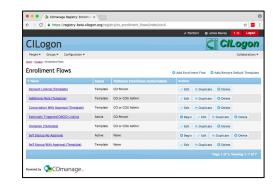
CILogon: 10+ year sustained effort to enable secure logon to scientific cyberinfrastructure (CI)

for seamless identity and access management (IAM) using federated identities (SAML, OIDC, OAuth, JWT, X.509, LDAP, SSH, etc.) so researchers log on with their existing credentials from their home organization supporting 17,500+ active users from 450+

with onboarding/offboarding/attributes/groups/roles managed consistently across multiple applications

organizations around the world

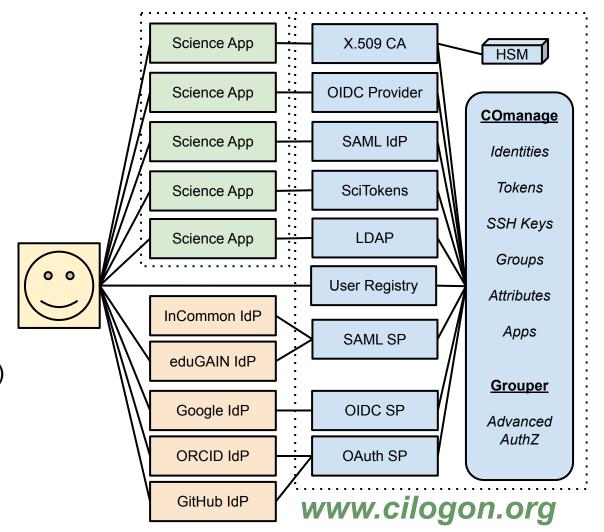






supporting access to science applications on HPC clusters, in Jupyter notebooks, using Globus, via REST APIs, and many other interfaces

using existing identity
providers from the
researcher's home
organization (SAML/ADFS)
or external sources
(Google, ORCID, GitHub)





realizing our vision

```
align with InCommon Trusted Access Platform
   (https://www.incommon.org/trusted-access/)
   Shibboleth, COmanage, Grouper
provide hosted services
   common IAM platform across many collaborations
   growing CILogon operations (since 2010)
   reliability / sustainability
```



Open Source

```
CILogon (https://github.com/cilogon)
   OpenID Connect, OAuth, X.509
InCommon (https://www.incommon.org/trusted-access/)
   Shibboleth, COmanage, Grouper
IdentityPython (https://idpy.org/)
    pyFF, SATOSA
SciTokens (https://scitokens.org/)
OpenLDAP with voPerson (https://voperson.org)
```



sustainability

development supported by NSF/DOE operational support from XSEDE



non-profit subscription model administered by NCSA/UIUC supports long-term sustainability

provides contracted SLAs

CILogon remains open source and focused on research & scholarship needs

https://www.cilogon.org/subscribe

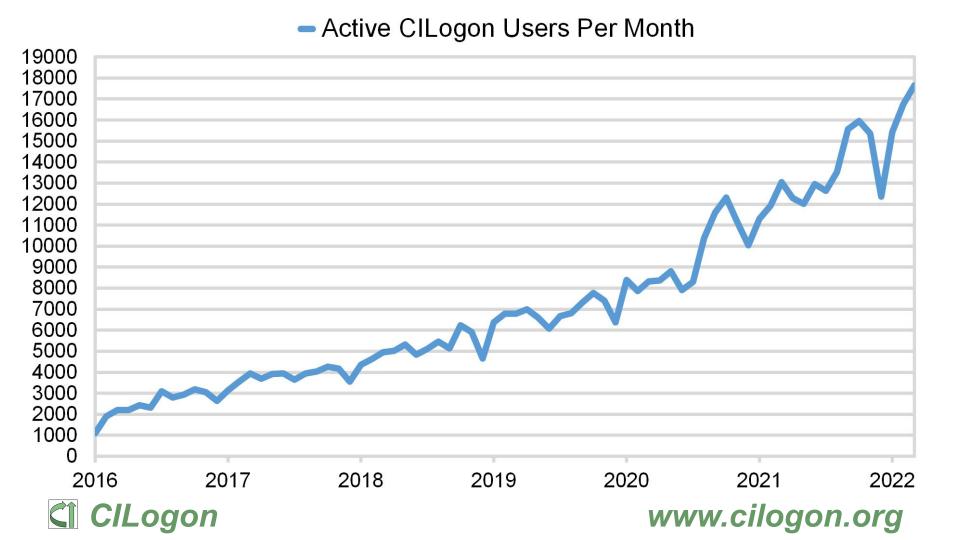


www.cilogon.org

our 10+ year history

- 2009 Federated login to TeraGrid. NSF ARRA award.
- 2010 CILogon operations begin. IGTF X.509 CAs operational.
- 2011 NSF SDCI award. OAuth support. InCommon Silver support.
- 2012 DOE ASCR award. Globus identity linking. InCommon R&S.
- 2013 XSEDE operations support. LIGO Data Grid use.
- 2016 NSF CICI award. eduGAIN support. OIDC support.
- 2017 COmanage support. AWS deployment.
- 2019 Transition to subscription funding model.
- 2020 Grouper and SATOSA support.
- 2021 InCommon Catalyst Program. SciTokens and WLCG JWT support.





Top 20 IdPs

(by # of unique active users in March 2022)

1315 Penn State	325 MIT
725 XSEDE	302 University of Chicago
720 University of Illinois at Urbana-Champaign	297 Washington University in St. Louis
665 Fermi National Accelerator Laboratory	264 NCSA
619 National Institutes of Health	255 Stanford University
514 LIGO Scientific Collaboration	251 Purdue University Main Campus
486 Northeastern University	240 University of Wisconsin-Madison
483 University of Michigan	216 Northwestern University
423 University of California-Los Angeles	208 University of California-San Diego
420 Michigan State University	189 Yale University



campus & researcher IDs

4,000+ identity providers available via eduGAIN including CERN, NCSA, LIGO, XSEDE, ...
OAuth-based identity providers
ORCID GitHub Google Microsoft supporting researcher mobility supporting researchers w/o campus IdPs



managing 1000s of IdPs

SAML Metadata Query Protocol

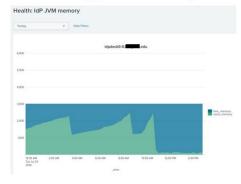
https://datatracker.ietf.org/doc/draft-young-md-query/https://spaces.at.internet2.edu/display/MDQ

monitor daily IdP changes

https://groups.google.com/a/cilogon.org/g/idp-updates https://cilogon.org/idplist/

Want to see how adopting the MDQ metadata service affects your system's performance?

Here is an in-use memory graph at the moment a university switched to using MDQ in its production IDP servers:



Great results from a Service Provider:

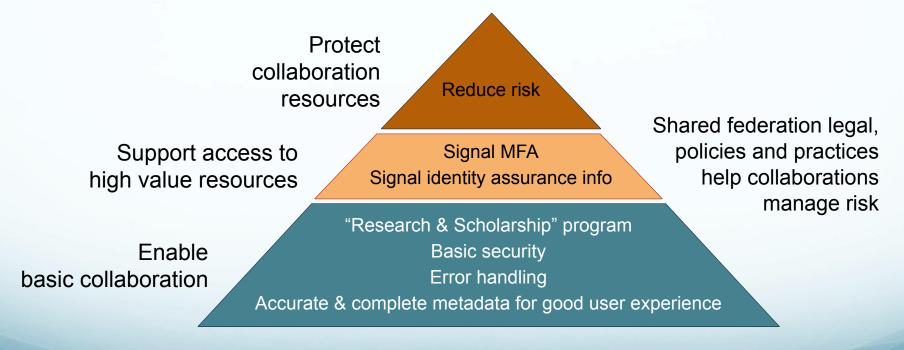
"We were able to switch to using MDQ. The service restarts in 5 seconds now versus 15 minutes."



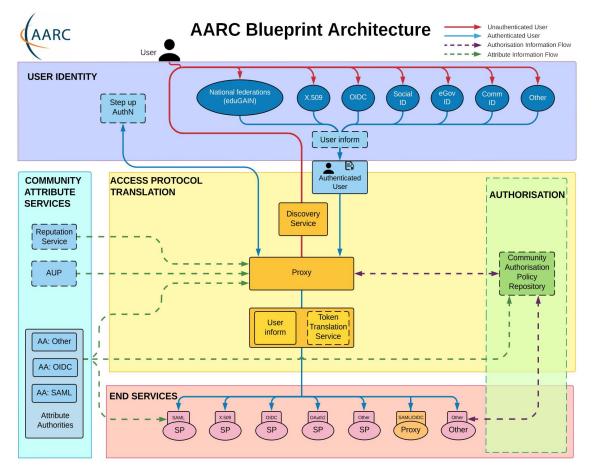




Federation support of research and scholarly communities



https://www.incommon.org/federation/baseline/



https://aarc-community.org/architecture/



federation proxy

apps don't need to handle the complexities of federation in isolation many apps can't handle 1000s of identity providers (e.g., AWS) a federation proxy service can handle federation for many (related) apps a federation proxy can handle targeted user identifiers consistently open source software for operating your own proxy:

SATOSA / SimpleSAMLphp

proxy as-a-service providers:

CILogon / eduTEAMS / Globus



tokens for science

OpenID Connect (OIDC) ID Tokens (e.g., SCiMMA) containing user attributes and group memberships from the research community (via COmanage) and from the researcher's home institution (via InCommon)





SciTokens (e.g., LIGO)

containing authorization scope values

determined by per client/subscriber policy



WLCG Tokens (e.g., Fermilab) support for wlcg.groups and storage.*|compute.* scopes



GA4GH Passports (e.g., Australian BioCommons) support for AffiliationAndRole, AcceptedTermsAndPolicies, ResearcherStatus, ControlledAccessGrants, and LinkedIdentities



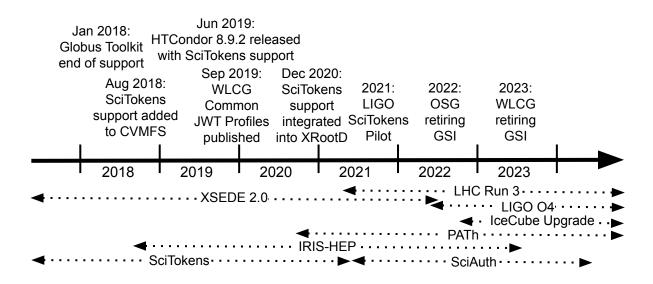
www.cilogon.org

token standards

- RFC 6749: OAuth 2.0 Authorization Framework
 - token request, consent, refresh
- RFC 7519: JSON Web Token (JWT)
 - · self-describing tokens, distributed validation
- RFC 8414: OAuth 2.0 Authorization Server Metadata
 - token signing keys, policies, endpoint URLs
- RFC 8693: OAuth 2.0 Token Exchange
 - token delegation, drop privileges (reduce "scope")
- . RFC 9068: JWT Profile for OAuth 2.0 Access Tokens
 - authorization claims using JWT "scope" and "aud"



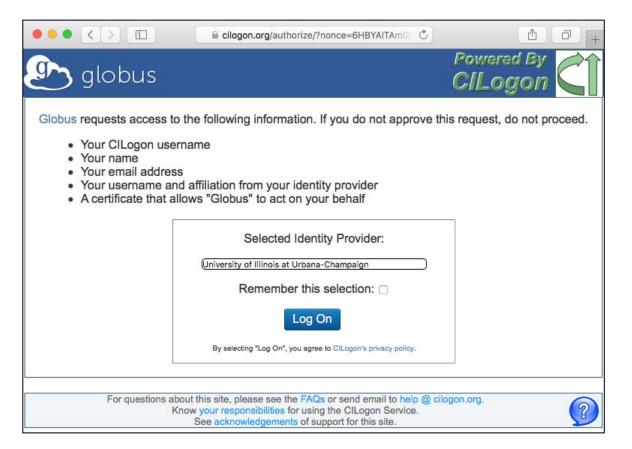
token transition timeline



https://opensciencegrid.org/docs/security/tokens/overview/ https://sciauth.org/



informed consent



examples of CILogon-enabled sites

2i2c, Apache Airavata Test Drive, Ask.Cl, ATLAS Connect, Australian BioCommons, BNL Quantum Astrometry, Brainlife.io, CADRE, CERN PanDA, Chem Compute, ClassTranscribe, CloudBank, Clowder, CMS Connect, Connect.ci, Custos, CyberGISX, CyVerse, DataCite, Duke CI Connect, Einstein Toolkit, FABRIC, Fermilab, Flywheel, GeoChemSim, Globus, GW-Astronomy, HubICL, HTRC, ImPACT, LIGO, LROSE, LS-CAT, LSST, Mass Open Cloud, MIT Engaging OnDemand, MSU HPCC OnDemand, MyGeoHub, NEON, NIH ClinOmics, NIH KnowEnG, Ocean Observatories Initiative, Open Science Chain, OSC OnDemand, OSG Connect, Pacific Research Platform, QUBES, SciGaP, SCiMMA, SEAGrid, SeedMeLab, SimVascular, Social Media Macroscope, UCLA JupyterHub, Vanderbilt JupyterHub, and XSEDE





"Managed Services to Simplify Cloud Access for Computer Science Research and Education"

University of California, San Diego (UCSD)'s San Diego Supercomputer Center (SDSC) and Information Technology Services (ITS) Division University of Washington (UW)'s eScience Institute University of California, Berkeley (UCB)'s Division of Data Science

https://www.cloudbank.org/





Partners

HOME / ABOUT

CloudBank has engaged with these important public partners. Click on the company logo to learn more.





Beam





Google Cloud









Log in

To **manage or access your CloudBank funds**, click the "Log in with <u>CILogon</u>" button using the instructions below. Login is currently limited to <u>funded</u> users. Contact <u>help@cloudbank.org</u> with any issues.

LOG IN WITH CILOGON

- 1. Click the "Log in with CILogon" button to visit a page for selecting login credentials.
- 2. Select an organization (Identity Provider) from the list to which you belong, and for which the email address matches your existing cloudbank.org account.
- 3. Click the "Log On" button to visit your organization's login page.
- 4. Login with your organization's credentials and you will be redirected back to cloudbank.org.







Consent to Attribute Release



<u>CloudBank</u> requests access to the following information. If you do not approve this request, do not proceed.

- Your ClLogon user identifier
- Your name
- · Your email address
- Your username and affiliation from your identity provider

Select an Identity Provider University of Illinois at Urbana-Champaign • Remember this selection Log On By selecting "Log On", you agree to ClLogon's privacy policy.

For questions about this site, please see the <u>FAQs</u> or send email to <u>help@cilogon.org</u>.

Know <u>your responsibilities</u> for using the CILogon Service.

See <u>acknowledgements</u> of support for this site.





Dashboard



Manage CloudBank Funds

View and manage your CloudBank funds and grant access to associated billing accounts (IAM).



Access CloudBank Billing Accounts

View all of your cloud billing accounts and current spend. Access public cloud web portals to manage cloud services.



Monitor & Optimize Your Usage

Access Nutanix Beam to monitor usage for each of accounts and view recommendations to optimize your usage.





Billing Account Access

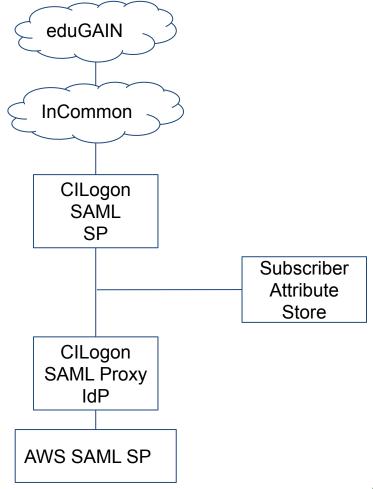
HOME / DASHBOARD

You have access to the following <u>billing accounts</u>. A billing account can be dedicated or shared with multiple people. If you are in a shared billing account, please note that we can only track usage at the tag and service level so consult your billing account manager for assistance if you need to track usage at a finer granularity.

BILLING ACCOUNT ID	CLOUD USERNAME	INITIAL PRIVILEGES GRANTED	PUBLIC CLOUD	PUBLIC CLOUD WEB CONSOLE LOGIN	FUND	AWARD AMOUNT	BALANCE
713660817510	skoranda	Login	Amazon Web Services	login → 3	Test Fund 1	\$51.00 *	\$41.39 (81%) *
test-2-1268	skoranda	Login (viewer)	Google Cloud Platform	login → 3	Test Fund 2	\$100.00 *	\$31.58 (32%) *

^{*} No specific limit specified on this *billing account*, value is inherited from *fund*.







www.cilogon.org

Thanks!

contact:

help@cilogon.org

subscribe for updates:

https://groups.google.com/a/cilogon.org/g/announce



extra slides



building blocks

InCommon Federation:

single sign-on for US R&E

eduGAIN:

global interfederation

REFEDS:

international standards for R&E federations

InCommon Trusted Access Platform:

open source IAM software



InCommon Trusted Access Platform

Shibboleth: federated single sign-on

COmanage: collaborative organization management

Grouper: enterprise group and access management (including

point-in-time auditing)

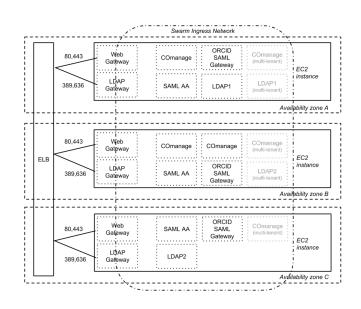
Midpoint: provisioning engine

https://www.incommon.org/trusted-access/



deployed to AWS

benefits of Net+ AWS
multiple availability zones
Docker containers in swarm mode
using R53 EC2 RDS ELB EFS





our baseline: REFEDS R&S

Attribute release continues to be the #1 stumbling block for new users.

We operate under the REFEDS R&S policy.

Does your campus support REFEDS R&S?

https://refeds.org/research-and-scholarship

https://cilogon.org/testidp/





security for global interfederation

We operate under the SIRTFI

framework: Security Incident Response

Trust Framework for Federated Identity

https://refeds.org/sirtfi

Supported by the NCSA Incident Response Team

https://security.ncsa.illinois.edu/







certificates for int'l science

CILogon CA policy update for int'l use approved in 2016 by Interoperable Global Trust Federation Requiring R&S + Sirtfi

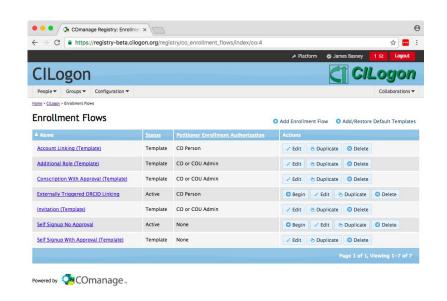






managing project groups/roles

COmanage provides: enrollment flows expiration policies self service permissions pipelines



https://www.cilogon.org/comanage





collaboration management platform built for federated identity

Open Source

Internet2/InCommon

PHP

20+ deployments managing more than 50K federated identities

OpenID Connect (OIDC)

third gen OpenID (after OpenID 1.0/2.0)

specifications: https://openid.net/connect/

authentication layer on top of OAuth 2.0 authorization framework (RFC 6749)

adds new token type: ID Token

adds new OAuth resource: UserInfo

standard claims and scope values





science gateways

enable web-based computational experiments and data management

CILogon-enabled hosted gateways:

Science Gateway Platform as a Service









nDemand

Supercomputing. Seamlessly. Open, Interactive HPC Via the Web

Support for Apache HTTP authentication modules including ADFS, CAS, OIDC, Shibboleth, Keycloak

CILogon support via Keycloak

Mapping federated identities to local accounts

https://openondemand.org/

https://osc.github.io/ood-documentation/latest/authentication





Step 2. Login via your provider
For example, here I've chosen Ohio State University as my provider and am



Step 3. Map it to your HPC account (first login only)
If it is the first time logging in with this provider, you will need to associate it with your
HPC account

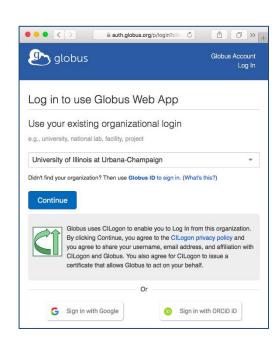


Log in with third party through ClLogon

Globus

campus authentication to your Globus Data Transfer Node

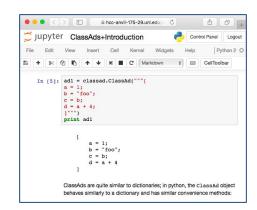
campus identities for Globus Auth





JupyterHub

notebooks support authoring/sharing of code, math, text, and multimedia federated authentication using CILogon one IdP or many





https://jupyterhub.readthedocs.io/en/latest/reference/authenticators.html https://github.com/jupyterhub/oauthenticator

https://zero-to-jupyterhub.readthedocs.io/en/latest/authentication.html



www.cilogon.org

Kubernetes

using Kubernetes native OIDC support public clients, refresh tokens, API access https://kubernetes.io/docs/reference/access-authn-authz/

demonstrated by PRP@UCSD

https://ucsd-prp.gitlab.io/userdocs/start/get-access/



Get access

BACK TO DDD



MediaWiki



custom COmanage user identifier assignment for MediaWiki username

MediaWiki's OIDC extension for auth

CILogon OIDC Provider sends custom MediaWiki username as sub claim

MediaWiki's OAuth extension for COmanage account provisioner

http://www.cilogon.org/mediawiki

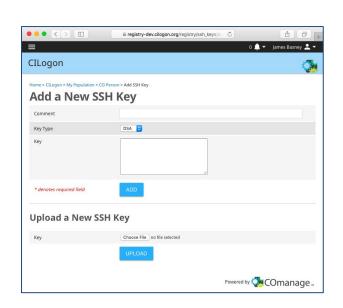


federated SSH keys

users register SSH public key during enrollment

associated with their federated identity provisioned to LDAP

used by SSH server for authorization



https://serverfault.com/questions/653792/ssh-key-authentication-using-ldap



support for idphint specification

https://www.cilogon.org/oidc

https://aarc-project.eu/guidelines/aarc-g049/

AARC-G049 A specification for IdP hinting

This document defines a generic browser-based protocol for conveying – to services – hints about the IdPs or IdP-SP-proxies that should be used for authenticating the principal.



attribute-based authz example

eduPersonAffiliation: specifies the person's relationship(s) to the institution in broad categories

permissible values: faculty, student, staff, alum, member, affiliate, employee, library-walk-in

specification: https://refeds.org/eduperson

use cases:

- software licenses
- data access restrictions
- resource allocation limits



role-based authz example: AWS

COmanage assigns Roles to each Person

via enrollment, approval workflows, expiration, etc.

linked to federated identities of the Person

CILogon SAML Proxy asserts Roles to AWS at authentication time

SAML proxy allows use of multiple identity providers

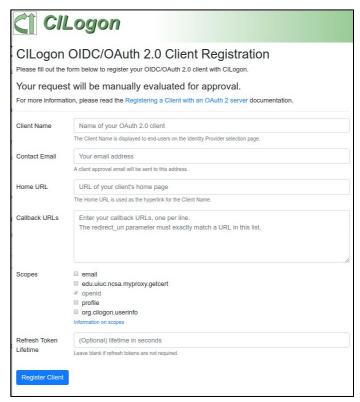
AWS IAM maps Roles to Permissions for access to AWS services

AWS Security Token Service (STS) provides temporary security credentials for CLI/API access



registering your OIDC app

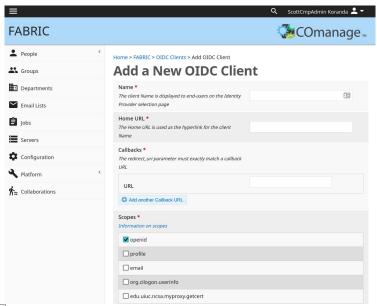
submit request at https://cilogon.org/oauth2/register including app details save client id and client secret wait for notification by help@cilogon.org see docs: http://www.cilogon.org/oidc

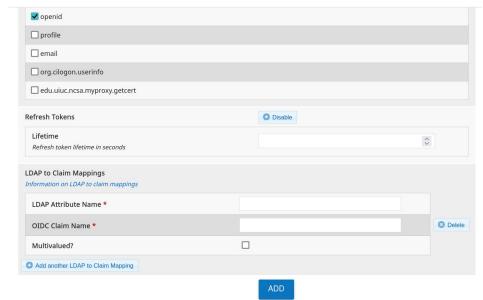




managing your OIDC apps

subscribers manage apps using COmanage







OAuth APIs for managing apps

subscribers can also manage OIDC apps via standard APIs:

- RFC 7591 OAuth 2.0 Dynamic Client Registration Protocol
- RFC 7592 OAuth 2.0 Dynamic Client Registration Management Protocol



configuring your OIDC app

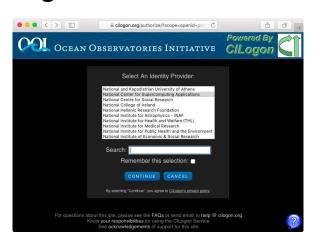
OIDC Discovery URL provides metadata

https://cilogon.org/.well-known/openid-configuration

contact help@cilogon.org to customize IdPs, claims, etc.

docs / examples:

http://www.cilogon.org/oidc





seamless campus integration

bypass CILogon screens when accessing local campus research applications

consent managed locally by campus

always use campus IdP

an OpenID Connect proxy to your campus SAML IdP

example: https://cybergateway.iu.edu/



bridging campus and VO IAM

passing campus and VO attributes to the application obtaining user consent via OIDC manage VO attributes in COmanage customize attributes/claims per app application-specific identifiers linking campus, researcher, and VO IDs driving authorization via group memberships



voPerson

an LDAP attribute schema (object class) with usage recommendations for VOs

voPersonApplicationUID	voPersonExternalID
voPersonAuthorName	voPersonID
voPersonCertificateDN	voPersonSoRID
voPersonCertificateIssuerDN	voPersonStatus

https://voperson.org/





"To actively support life science research communities with community scale digital infrastructure developed and maintained in concert with international peer infrastructures."

Target ~30,000 life science researchers in AU





"Established in 2009, the Australian Access Federation (AAF) is Australia's identity federation and part of a global network of over 61 federations around the world."



Initial Collaboration Target



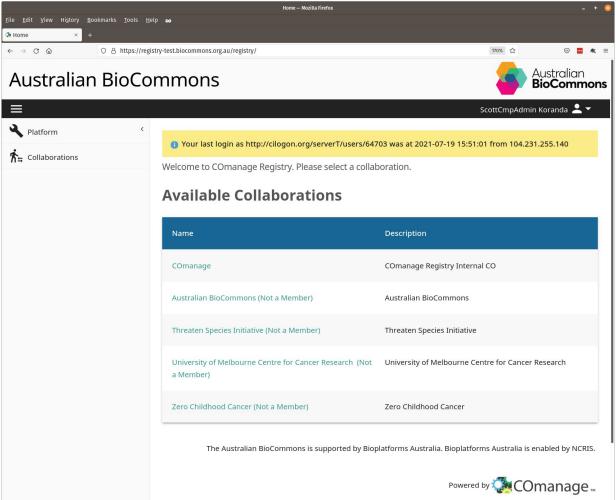


University of Melbourne Centre for Cancer Research (UMCCR) "Driving innovation and implementation for clinical impact in cancer care"

Zero Childhood Cancer Program (ZERO)

"Australia's first-ever personalised medicine program for children and young people with high-risk cancer"







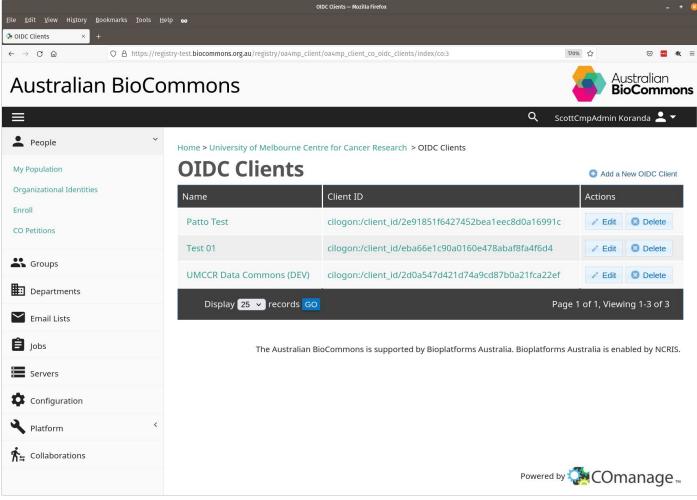




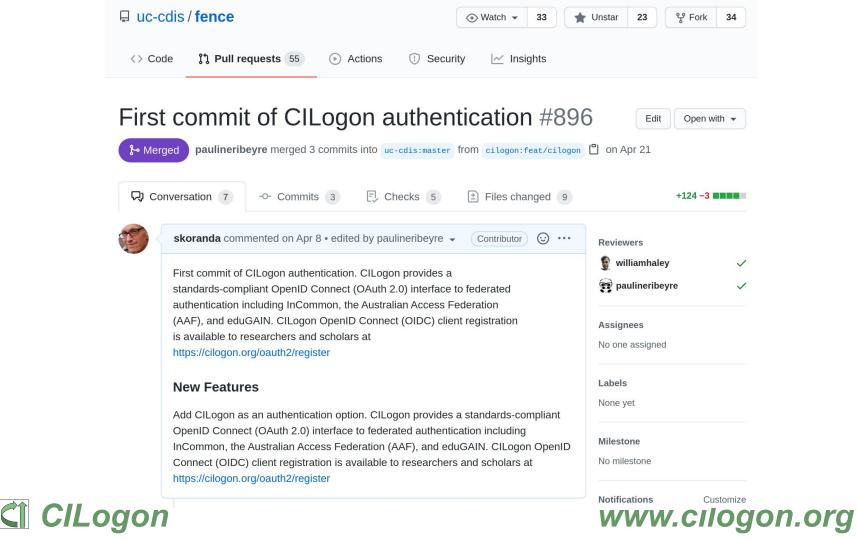
Center for Translational Data Science, University of Chicago

"Gen3 Data Commons are cyberinfrastructure that co-locates data analysis, exploration and visualization tools with data management services for import and export of structured information like clinical, phenotypic, or biospecimen data, and data objects, like genomics data files or medical images."









Global Authz Interoperability



GA4GH Passports and the Authorization and Authentication Infrastructure



GA4GH AAI OpenID Connect Profile

"In particular, this specification introduces a JSON Web Token (JWT) syntax for an access token to enable an OIDC provider (called a Broker) to allow a downstream access token consumer (called a Claim Clearinghouse) to locate the Broker's /userinfo endpoint as a means to fetch GA4GH Claims. This specification is suggested to be used together with others that specify the syntax and semantics of the GA4GH Claims exchanged."





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!3fsruagzH47Z800fjwaXGRnFVR8=",
   "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": [
    "eyJ0eXAi0iJKVTQiLCJraWQi0iIyNDRCMjM1RjZCMjhFMzQxMDhEMTAxRUFDNzM2MkM0RSIsImFsZyI6IlJTMjU2In0.eyJzdWIi0iJodHRw0i8vY2lsb2dvbi5vcmcvc2Vyd
    "eyJ0eXAi0iJKV1QiLCJraWQi0iIyNDRCMjM1RjZCMjhFMzQxMDhEMTAxRUFDNzM2MkM0RSIsImFsZyI6IlJTMjU2In0.eyJzdWIi0iJodHRw0i8vY2lsb2dvbi5vcmcvc2Vyd
    evJ0eXAiOiJKV10iLCJraW0iOiIvNDRCMiM1RiZCMihFMz0xMDhEMTAxRUFDNzM2MkM0RSIsImFsZvI6IlJTMiU2In0.evJzdWIi0iJodHRw0i8vY2lsb2dvbi5vcmcvc2Vvd"
   "email": "skoranda@illinois.edu",
   "cid": "cilogon:test.cilogon.org/demo0"

    Show/Hide Access Token
```

- Show/Hide ID Token
- Show/Hide certificate subject

