

# Authorizing Access to Science Gateway Resources

Jim Basney (NCSA & Trusted CI)

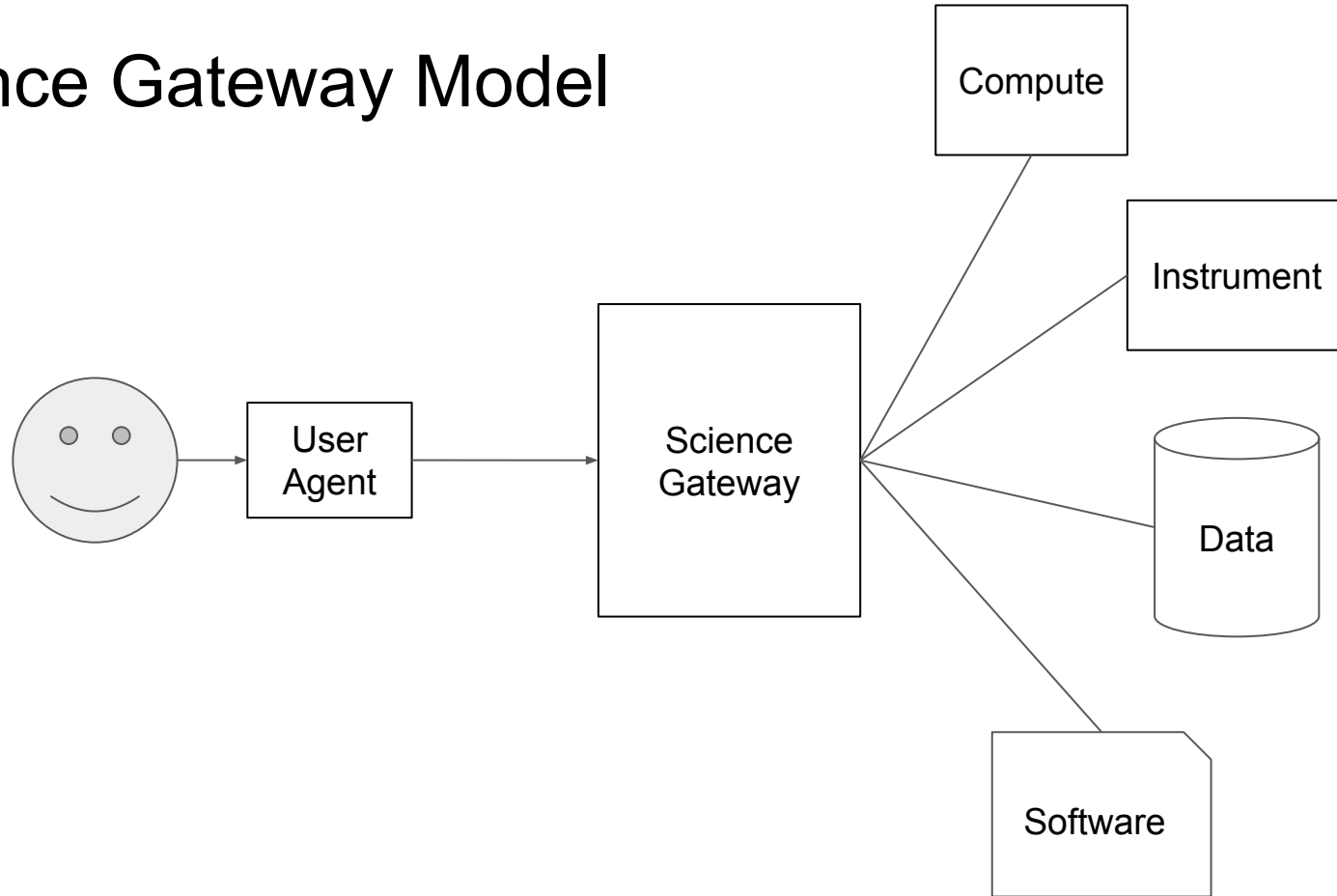
Marlon Pierce (Indiana University & SGCI)

Tom Barton (University of Chicago & Internet2)

<https://sciencegateways.org/engage/webinars>

Jan 9 2019

# Science Gateway Model



# Authorization

## Policies

- Acceptable use
- Resource limits
- Restricted-access scientific instruments
- Pre-publication research collaborations
- Data use agreements
- Controlled-access data sets

## Procedures

- User affiliation
- Self-organizing collaboration groups
- BYO resources
- Peer review allocations
- Blacklisting

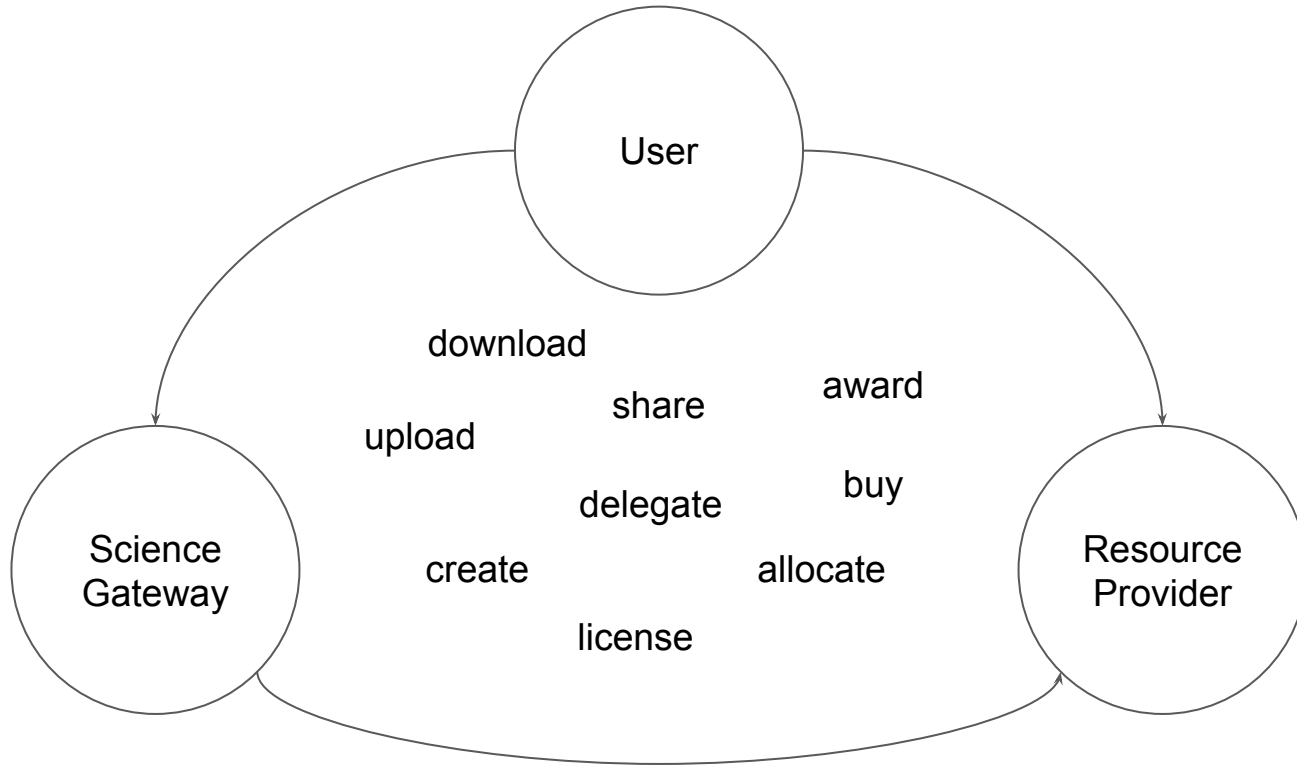
## Mechanisms

- User Attributes
- Groups
- Roles
- Allocations / Quotas
- Delegated Management

# Authorization

- Identity-based
  - User identifiers and access control lists
- Attribute-based
  - Access policies based on user attributes
- Role-based
  - Access controls based on group memberships and roles
- Capability-based
  - Tokens allows actions on resources

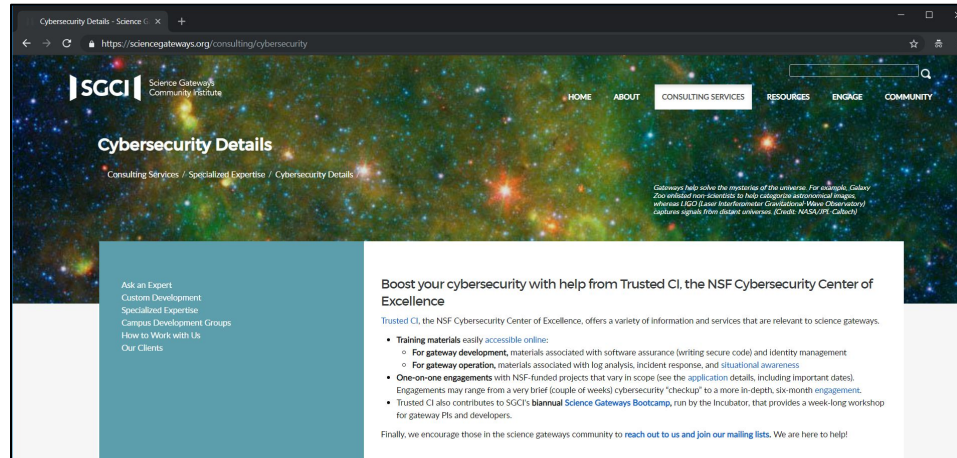
# Who owns the resource?



# SGCI and Trusted CI



- Trusted CI offers specialized engagements, or consultations, to science gateway developers and operators seeking cybersecurity support
- Trusted CI's partnership with SGCI includes training bootcamps, webinars, and direct support
- <https://trustedci.org/sgci/>



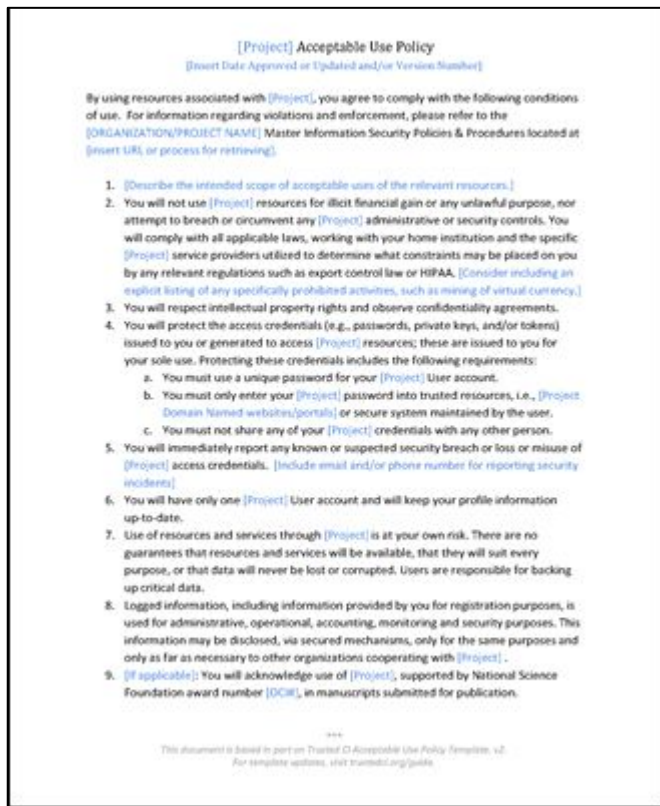
# Acceptable Use Policy

AUP can play an important role in your authorization approach:

- Communicate expectations to users
- Document consequences for violating policy
- Can require explicit acknowledgement on signup and policy change

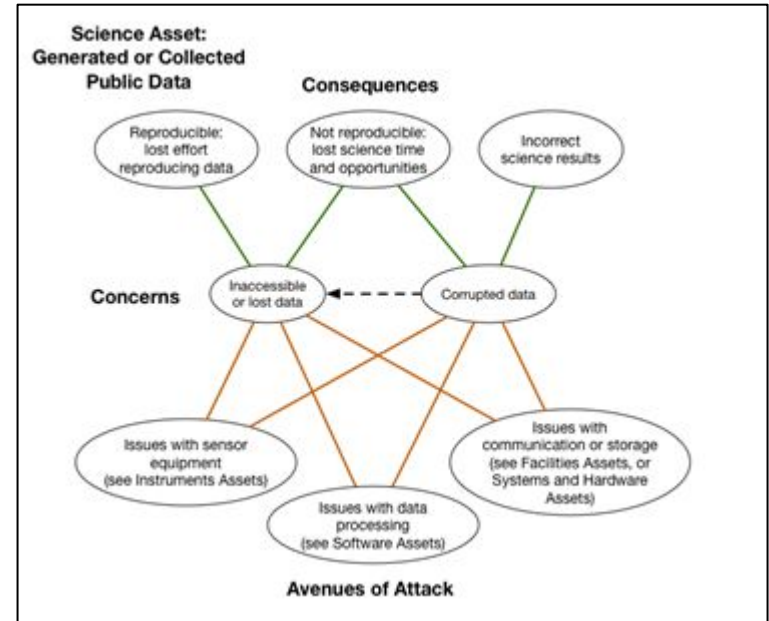
Example AUP from Trusted CI

- <https://trustedci.org/guide/>



# Open Science Cyber Risk Profile

- Provides an enumeration of common scientific assets and the IT risks associated with each
  - Scientific assets are resources critical to science mission
  - Focus on consequences to science mission rather than specific actors/tactics/vulnerabilities
  - List of common science assets. Each linked to a diagram showing science concerns, consequences, and avenues of attack.
- <https://trustedci.github.io/OSCRP>



# Delegated Authorization

- Transitive Mode
  - User is authorized to access Science Gateway
  - Science Gateway is authorized to access other resources
- Authorization Credentials Mode
  - Science Gateway accesses other resources via user-specific credentials
  - OAuth model: “An application making protected resource requests on behalf of the resource owner and with its authorization” (RFC 6749)

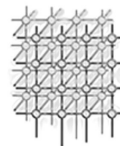
Von Welch, Jim Barlow, James Basney, Doru Marcusiu, Nancy Wilkins-Diehr, "A AAAA model to support science gateways with community accounts," *Concurrency and Computation: Practice and Experience*, Volume 19, Issue 6, March 2007.  
<https://doi.org/10.1002/cpe.1081>

Jim Basney, Von Welch, and Nancy Wilkins-Diehr, "TeraGrid Science Gateway AAAA Model: Implementation and Lessons Learned," *TeraGrid Conference*, August 2-5, 2010, Pittsburgh, PA. <https://doi.org/10.1145/1838574.1838576>

CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE  
*Concurrency Comput.: Pract. Exper.* 2007; 19:893–904  
Published online 10 October 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cpe.1081

## A AAAA model to support science gateways with community accounts

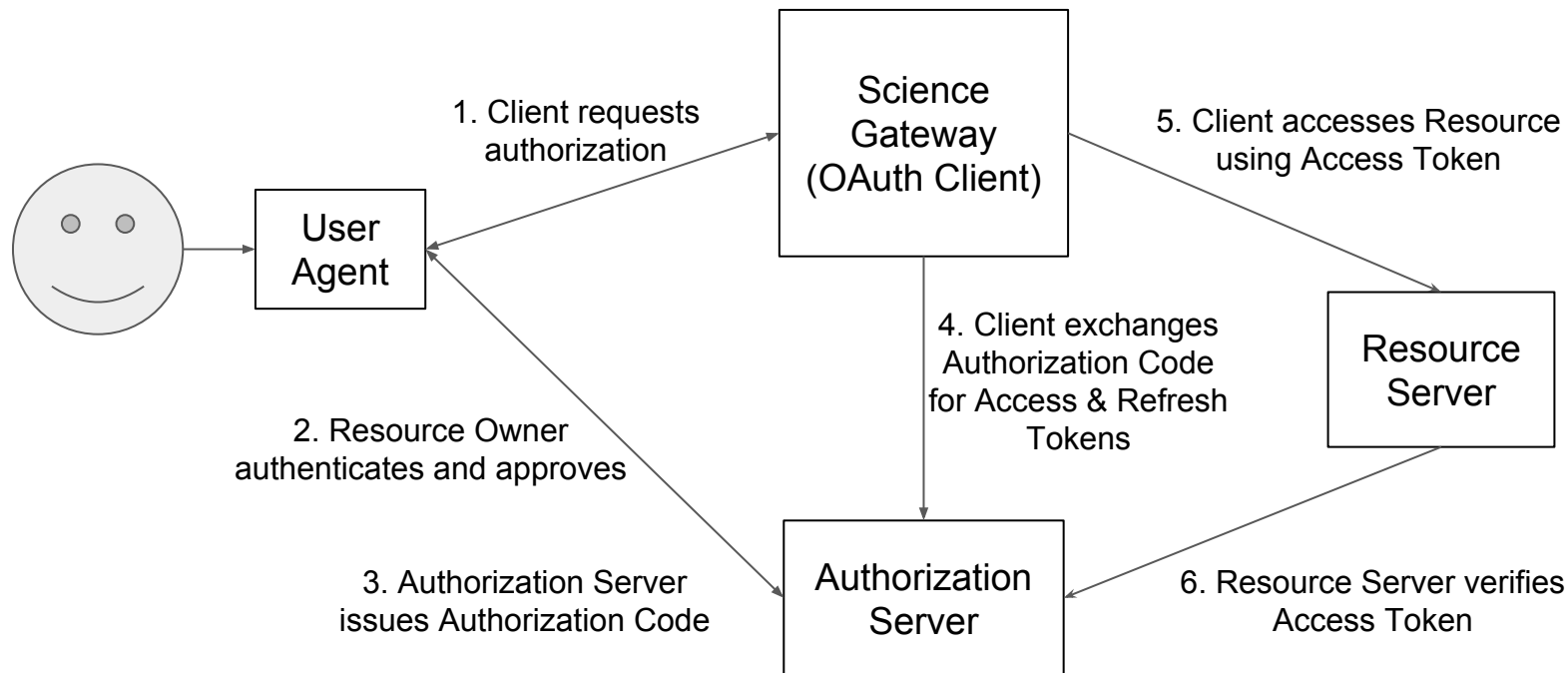
Von Welch<sup>1,\*</sup>, Jim Barlow<sup>1</sup>, James Basney<sup>1</sup>,  
Doru Marcusiu<sup>1</sup> and Nancy Wilkins-Diehr<sup>2</sup>



<sup>1</sup>National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign,  
1205 W. Clark Street, Room 1008, Urbana, IL 61801, U.S.A.

<sup>2</sup>San Diego Supercomputer Center (SDSC), University of California at San Diego, MC 0505,  
9500 Gilman Drive, La Jolla, CA 92093-0505, U.S.A.

# OAuth for Science Gateways



# OAuth: Scope and Consent



## Authorize Travis CI for Open Source



Travis CI for Open Source by [travis-ci](#)

wants to access your [jbasney](#) account



### Personal user data

Email addresses (read-only)



### Repository webhooks and services

Read and write access



### Commit statuses

Read and write access



### Deployments

Manage deployments and deployment status



### Organizations and teams

Read-only access



## Travis CI for Open Source

🕒 Last used within the last 2 weeks

👤 Developed by [travis-ci](#)

🔗 <https://travis-ci.org>

## Permissions






Revoke access

- ✓ Read org and team membership
- ✓ Access commit status
- ✓ Access deployment status
- ✓ Access user email addresses (read-only)
- ✓ Write repository hooks

Applications act on your behalf to access your data based on the permissions you grant them. Organizations control which applications are allowed to access their private data. Applications you authorize will always have access to public data in your organizations. [Read about third-party access.](#)

# SGCI - Internet2 Partnership

Just starting out - web site not even updated yet!

	Federated access
	Research networks
	R&E wifi roaming
	IAM/AAI software
	Commercial clouds*

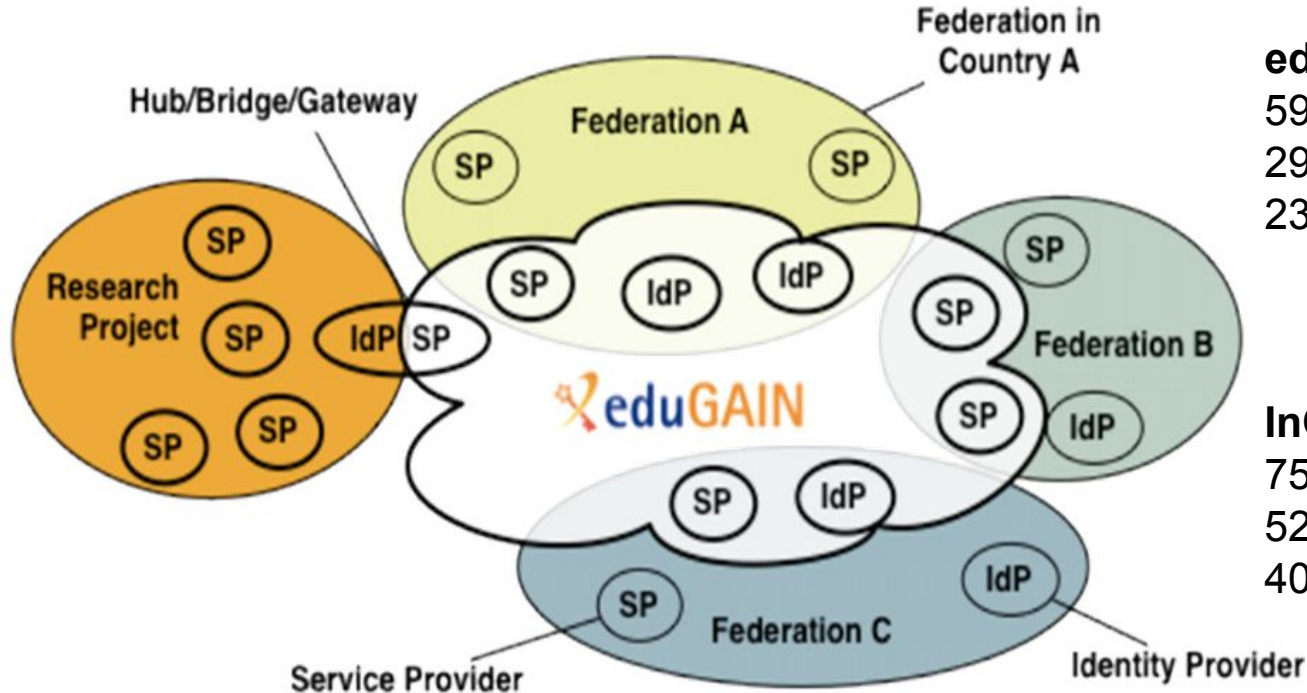
\* as experience is gained under [Exploring Clouds for Acceleration of Science](#) program

# Internet2 IAM Open Source Software

	Shibboleth - SAML federation, OIDC being added
	Grouper - enterprise grade access management
	COmanage – tailored for academic collaborations (Jim to say more later)
	MidPoint partnership - enterprise grade identity registry

More than just software packages - the real value lies in the Internet2 community's IAM expertise

# Federated Access - Global Reach



## **eduGAIN**

59 countries

2924 Identity Providers

2318 Service Providers

## **InCommon** (US Federation)

755 organizations

526 Identity Providers

4002 Service Providers

Each national R&E Federation decides which of their IdPs and SPs to publish into eduGAIN

# Federation Beyond Plain Authentication

## Value adds

- Research & Scholarship attributes (R&S program)
  - Name, email, persistent ID, affiliation
- MFA
- Assurance profiles
- Site logos, technical, security, & admin contacts
- Incident response procedure

## Pain points

- Sites that don't pay attention and Federations that don't manage that well
  - Low R&S adoption by IdPs
  - Hit/miss logos & contacts
    - Outside of InCommon
- Interop issues when different R&E Federations publish different views of eduGAIN

# Access Authorization Models and Examples

- Appendix B of “[Federated Identity Management for Research v2](#)” describes how 14 different research communities do it
  - A leading example is [ELIXIR](#), used as a gateway by several Life Sciences research communities
  - Most of these implement the [AARC Blueprint Architecture](#)
  - [CILogon](#) does too, as Jim will show
- 
- Related: [Globus High Assurance](#) for data transfer into and out of a secure/protected environment

# ELIXIR AAI Overview

## User authentication services

### Credential translation

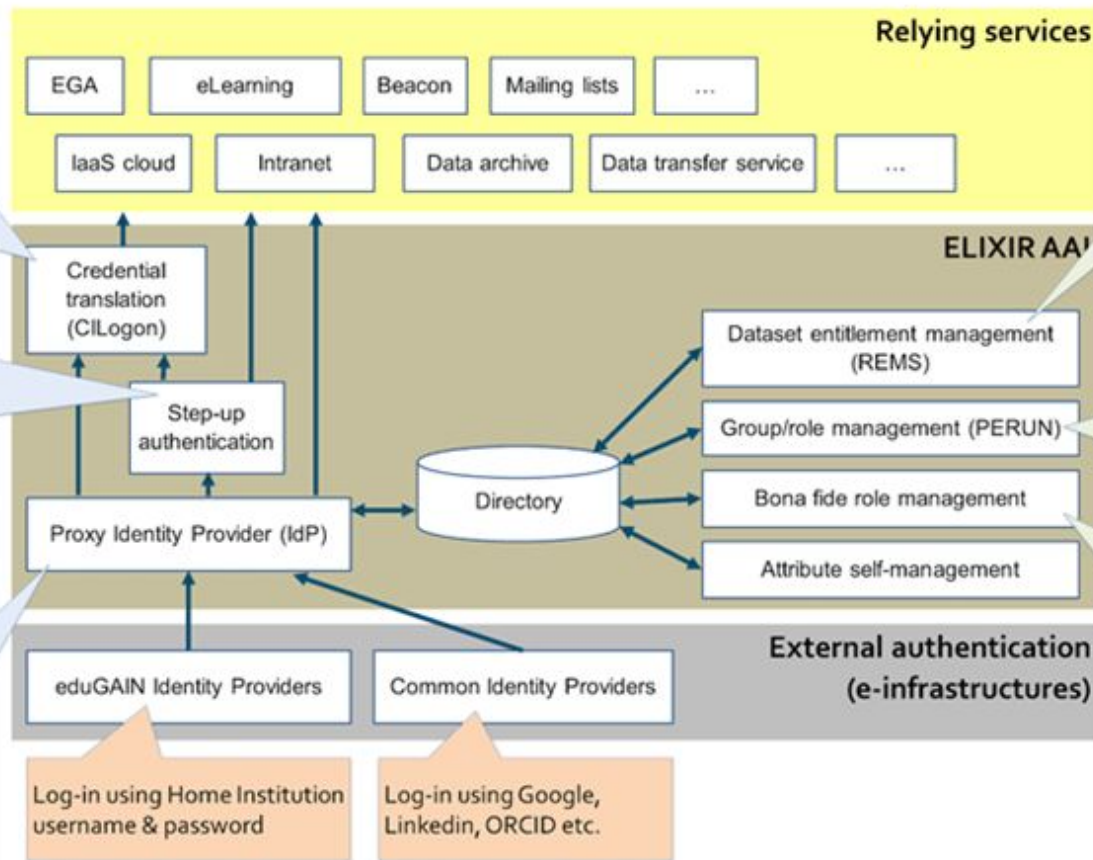
- Proxy IdP is WWW
- Some services are non-web e.g.
  - SSH access to a cloud VM
  - triggering file transfer between data archives
- Conversion to X.509

### Step-up Authentication

1. User authenticates weakly using external authentication
2. User authenticates with a second factor e.g. SMS-OTP or a mobile app

### Proxy Identity Provider

- User has one ELIXIR identity
- User can authenticate using external identities
- Proxy Identity Provider consolidates the IDs



## User authorisation services

### Dataset entitlements

- User applies for access rights to a dataset
- DAC approves → Endorsed user

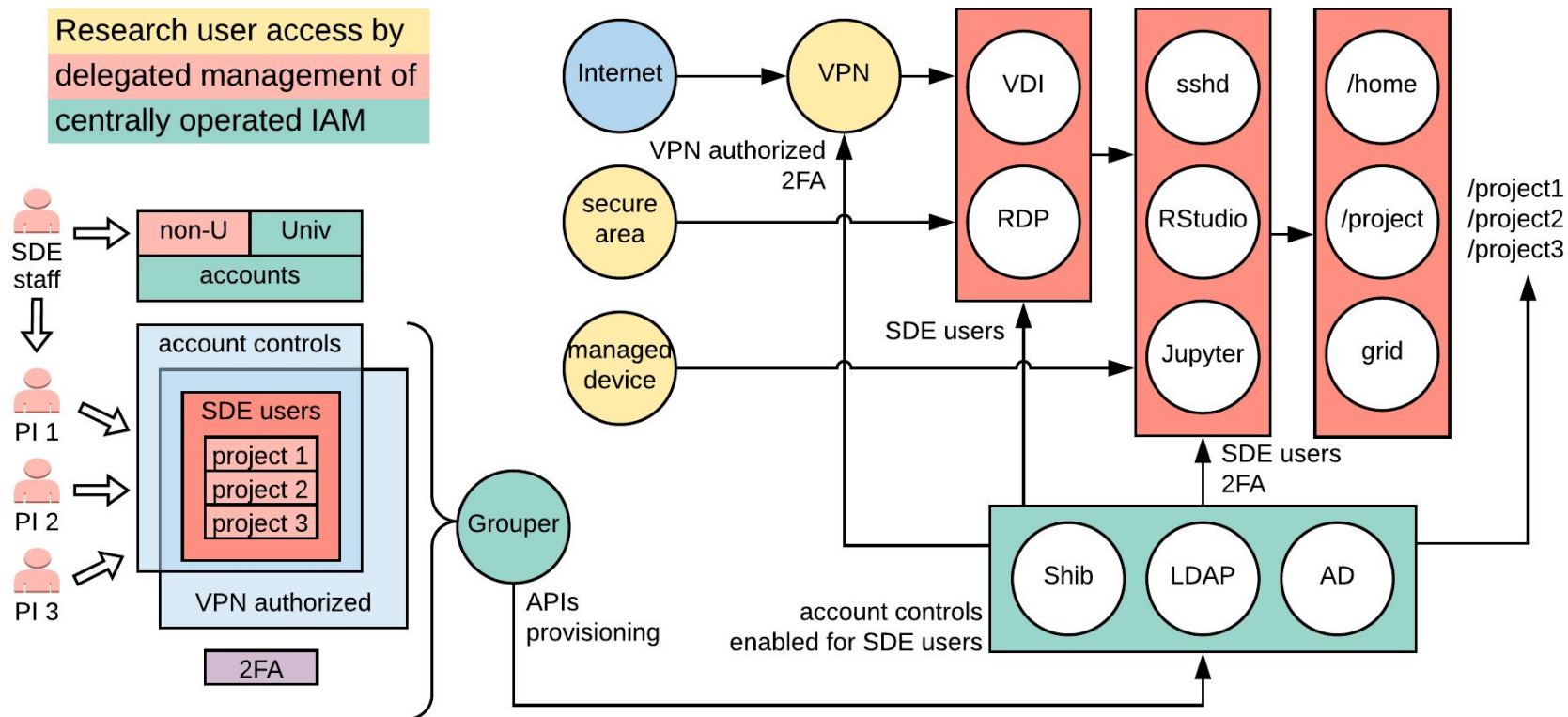
### Group/role management

- User can belong to one or more groups
- Group memberships and roles are managed by group managers

### Bona fide role management

- User applies for a bona fide status
- User commits to a Code of Conduct
- The community approves → Bona fide user

# UChicago Secure Data Enclave (NIST SP800-171)



# Science Gateway Security with Custos

Investigators: Marlon Pierce (IU), Suresh Marru (IU), Jim Basney (UIUC), Enis Afgan (JHU)

Senior Personnel: Vahid Jalili (OHSU), Jeff Gaynor (UIUC), Terry Fleury (UIUC), Marcus Christie (IU)

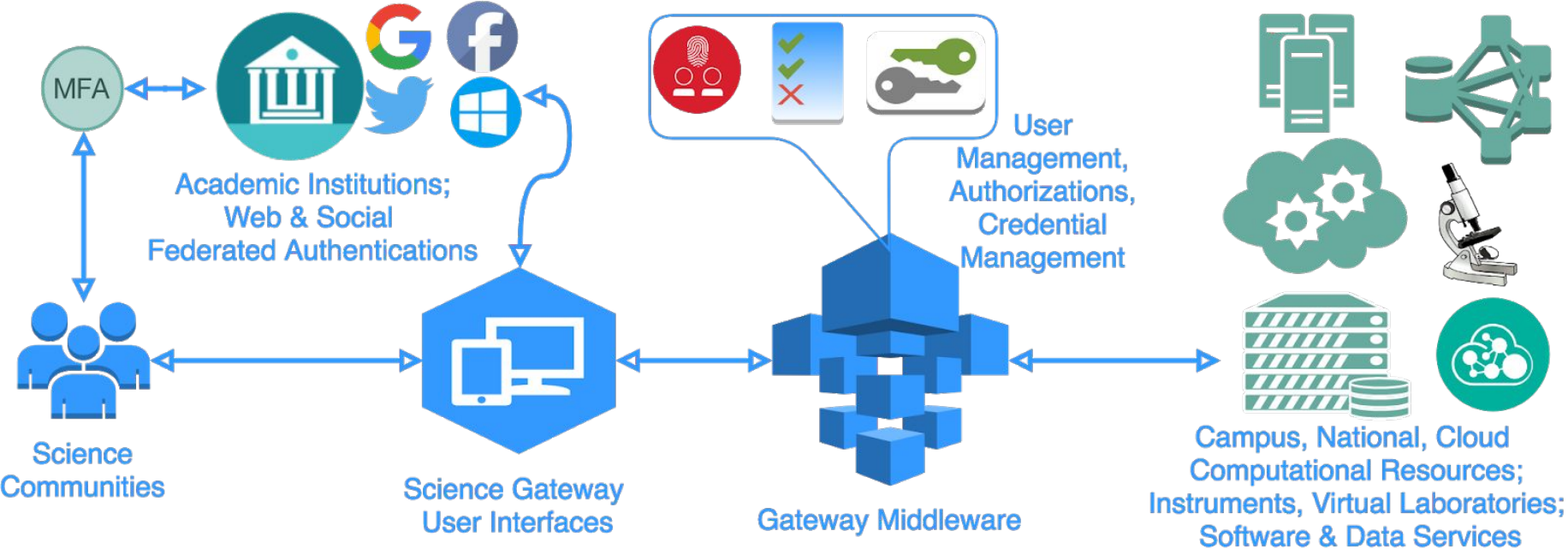
NSF CICI Award #1840003

# Science Gateways and Science

- Science gateways support scientific research for hundreds of thousands of users around the world.
- Gateways help scientists **produce, reproduce, manage, and publish** scientific digital content
- Pierce, Marlon E., Mark A. Miller, Emre H. Brookes, Mona Wong, Enis Afgan, Yan Liu, Sandra Gesing, Maytal Dahan, Suresh Marru, and Tony Walker. "Towards a Science Gateway Reference Architecture." (2018).
  - Proceedings of IWSG 2018
  - <https://scholarworks.iu.edu/dspace/handle/2022/22235>

# And so...

- Cybersecurity for science gateways needs to mature
- We need to consider the threats to all open digital science considered in the Open Science Cyber Risk Profile (OSCRP):
  - <http://trustedci.github.io/OSCRP/OSCRP.html>



Gateways need to manage three important aspects of security:

- User management: authentication, user profiles
- Resource connection management: credentials, keys, and security tokens for accessing third party resources
- Digital object management: manage the sharing of digital representations of experiments, resources, etc
  - Content created by the users and operators of a gateway

# What's At Stake?

Asset	Risk	Mitigation
User Identities	Unauthorized access to third party resources and content; password compromises.	Gateway account management should be based on general purpose identity management solutions, preferably based on transparent, standards-based, active and widely used open source software.
Third Party Resource Access	Gateways mismanage access to cyberinfrastructure such as XSEDE supercomputers, campus computing and storage, commercial cloud access.	Gateways should use software or services that specialize in managing secrets used for accessing remote resources. Gateways should likewise adopt best practices and protocols such as OAuth2 for interacting with remote services.
Gateway Content (Digital Objects)	Malicious deleting and alterations of content, inappropriate exposure of restricted content, and unauthorized use of restricted resources	Gateways should reuse best of breed, dedicated software or services for this feature rather than implementing it themselves.

Lost trust in gateways by users, resource providers, and the general scientific community

Gateways should move away from operating their own cybersecurity solutions for identity management, secrets management, and sharing

- These are tricky things to do well
- Operations, maintenance are hidden costs

Gateway cybersecurity should be based on open source software

- Community driven
- Implements best practices
- One source that everyone can inspect, audit, potentially contribute to

Gateway cybersecurity should be a service

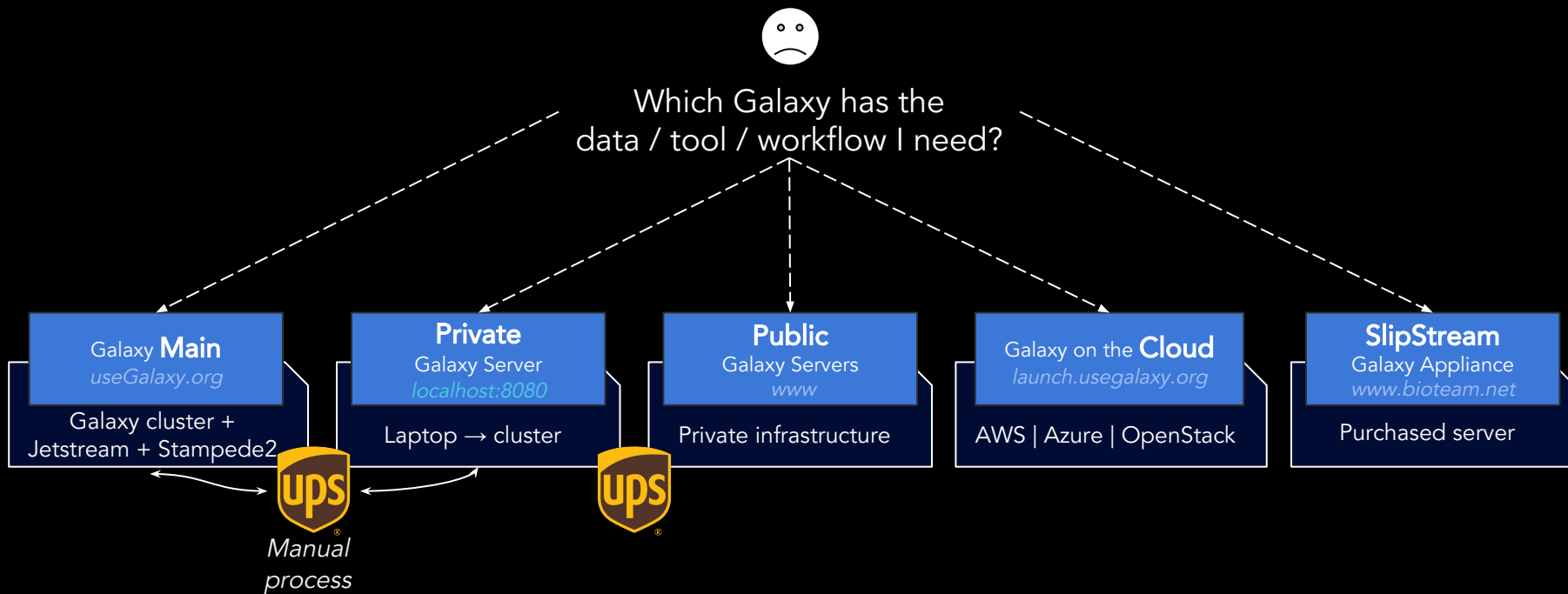
- Operate it using best cybersecurity practices
- Provide API-based access
- Provide open source “infrastructure as code” deployment, so the service is auditable just like the implementation

Custos's Founding Premises

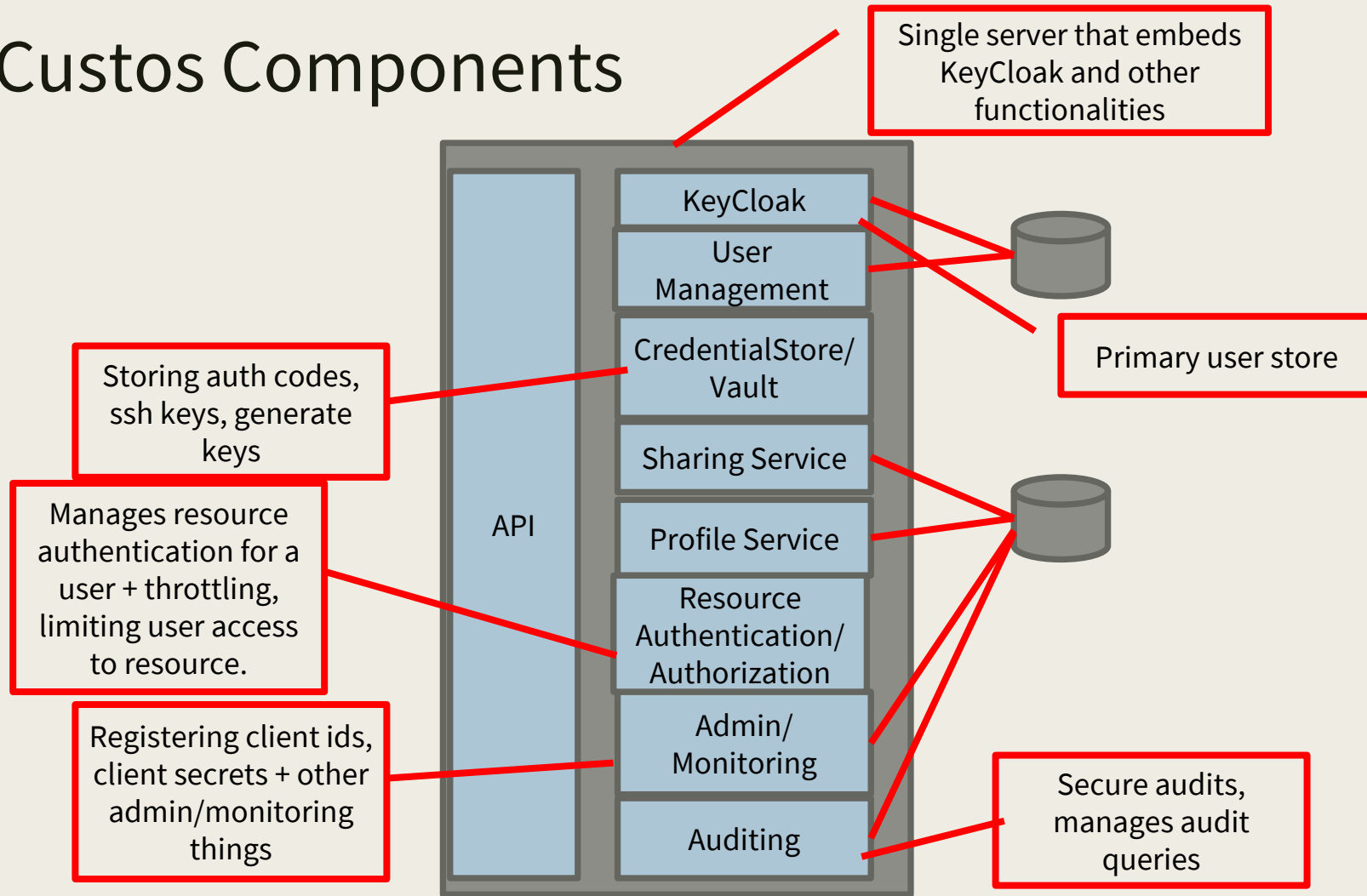
# Overall Plan for Custos

- Many Custos capabilities are currently implemented in Apache Airavata
  - We can create a tenant for you now that provides Custos core capabilities
- But we want to extract these capabilities as a new, standalone project
  - Support Galaxy and other non-Apache Airavata gateways
    - JupyterHub
  - Leverage CILogon, SciTokens, Galaxy expertise with cloud integration and scale
  - Take it to the Apache Software Foundation
- We need to take Custos through additional security reviews
  - Yearly, with Trusted CI
- We need to consider integrating third-party software such as HashiCorp's Vault

# Scaling challenges: **siloed login and fragmentation**



# Custos Components



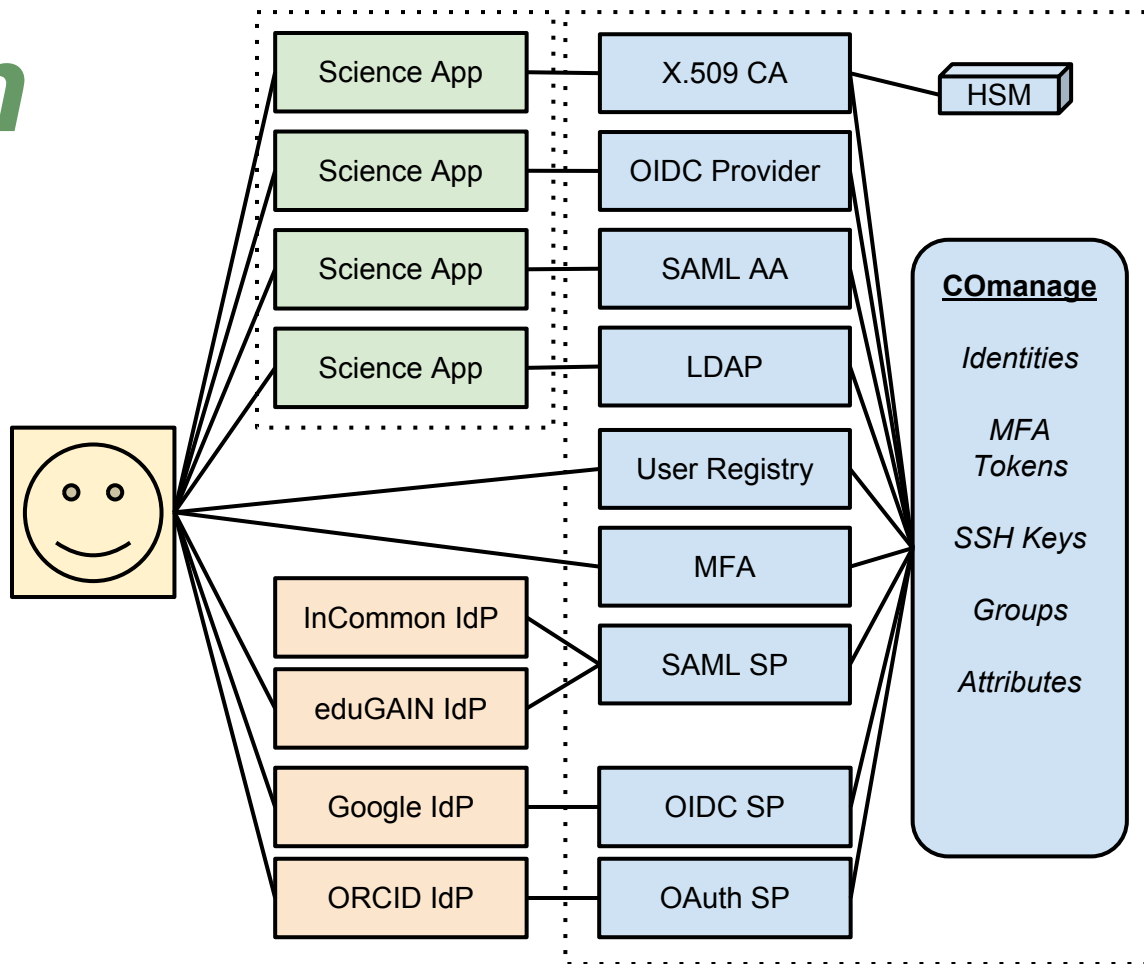
# Getting Involved

- Developer discussions will be on [dev@airvata.apache.org](mailto:dev@airvata.apache.org) until we are ready to make this a standalone project



# CILogon

an open source  
identity and  
access  
management  
platform for  
research  
collaborations



[www.cilogon.org](http://www.cilogon.org)

# eduPersonAffiliation: Campus Attribute for AuthZ

- 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: <http://macedir.org/specs/eduperson/#eduPersonAffiliation>
- Science Gateway use cases:
  - Software licenses
  - Data access restrictions
  - Resource allocation limits

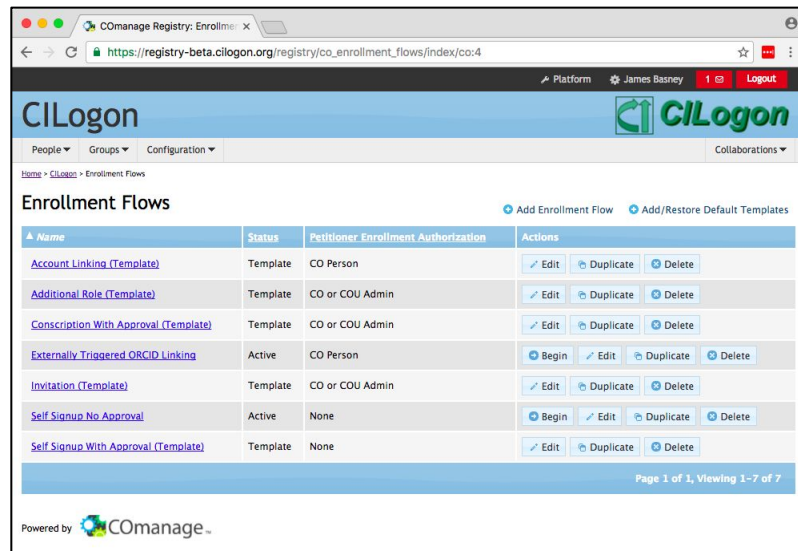
# eduPersonAffiliation: SAML example

```
<saml2:AttributeStatement>
  <saml2:Attribute FriendlyName="eduPersonScopedAffiliation"
    Name="urn:oid:1.3.6.1.4.1.5923.1.1.1.9"
    NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri">
    <saml2:AttributeValue>employee@illinois.edu</saml2:AttributeValue>
    <saml2:AttributeValue>member@illinois.edu</saml2:AttributeValue>
    <saml2:AttributeValue>staff@illinois.edu</saml2:AttributeValue>
  </saml2:Attribute>
  <saml2:Attribute FriendlyName="displayName" Name="urn:oid:2.16.840.1.113730.3.1.241"
    NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri">
    <saml2:AttributeValue xmlns:xsd="http://www.w3.org/2001/XMLSchema"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:type="xsd:string">James Alan Basney</saml2:AttributeValue>
  </saml2:Attribute>
  <saml2:Attribute FriendlyName="eduPersonPrincipalName" Name="urn:oid:1.3.6.1.4.1.5923.1.1.1.6"
    NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri">
    <saml2:AttributeValue xmlns:xsd="http://www.w3.org/2001/XMLSchema"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:type="xsd:string">jbasney@illinois.edu</saml2:AttributeValue>
  </saml2:Attribute>
</saml2:AttributeStatement>
```

# CILogon Groups/Roles (Powered by COnmanage)

Authorize based on group memberships and roles managed by:

- Custom enrollment flows
- Automated expiration policies
- Self service permissions
- Pipelines & Plugins for custom workflows and integrations (e.g., account provisioning)



<https://www.cilogon.org/comanage>

# CILogon ID Token example

```
{
  "sub": "jbasney@ncsa.illinois.edu",
  "eppn": "jbasney@ncsa.illinois.edu",
  "iss": "https://test.cilogon.org",
  "given_name": "James",
  "family_name": "Basney",
  "aud": "myproxy:oa4mp,2012:/client_id/180d79858441e8270aa6e199f9afaab8",
  "acr": "urn:oasis:names:tc:SAML:2.0:ac:classes:PasswordProtectedTransport",
  "idp": "https://idp.ncsa.illinois.edu/idp/shibboleth",
  "affiliation": "staff@ncsa.illinois.edu;employee@ncsa.illinois.edu;member@ncsa.illinois.edu",
  "uid": "jbasney",
  "uidNumber": "25555",
  "name": "James Basney",
  "isMemberOf": [ { "name": "lsst_users", "id": 1363 }, { "name": "lsst_int_lspdev", "id": 1618 } ],
  "email": "jbasney@illinois.edu",
  "exp": 1532630945,
  "iat": 1532630045,
  "auth_time": "1532630005"
}
```



- Open source software demonstrating *capabilities-based authorization* for distributed scientific computing
  - An alternative to identity-based or attribute-based authorization
  - Using CILogon, HTCondor, CVMFS, XRootD
- Using web standards
  - RFC 6749: OAuth 2.0 Authorization Framework
  - RFC 7519: JSON Web Token (JWT)
  - RFC 8414: OAuth 2.0 Authorization Server Metadata

[www.scitokens.org](http://www.scitokens.org)

# Example JSON Web Token



- The decoded token contains multiple scopes - basically filesystem authorizations.
- The audience narrows who the token is intended for.
- The issuer identifies who created the token; value used to locate the public keys needed to validate signature.
- The subject is an identifier for the resource owner.
- The expiration is a Unix timestamp when the token expires.

HEADER: ALGORITHM & TOKEN TYPE

```
{  
  "typ": "JWT",  
  "alg": "RS256"  
}
```

PAYLOAD: DATA

```
{  
  "scope": "read:/protected write:/store/u25321",  
  "aud": "https://demo.scitokens.org",  
  "iss": "https://demo.scitokens.org",  
  "sub": "bbockelm@cern.ch",  
  "exp": 1526954997,  
  "iat": 1526954397,  
  "nbf": 1526954397,  
  "jti": "78c44ce9-62bb-43e8-a7a6-f035f7ebd42b"  
}
```

# Summing Up

- Authorization policies, procedures, and mechanisms
- Authorization models and examples
- Identity-based, attribute-based, role-based, and capability-based authorization
- OAuth and JWT standards
- Good security and access management software is hard, and there are excellent open source options
  - Don't roll your own!
- There's lots of help available to gateway developers and operators
  - SGCI Partners like Trusted CI and Internet2
  - Campus central IT may have good IAM to leverage

Thanks!

Contact:

[jbasney@ncsa.illinois.edu](mailto:jbasney@ncsa.illinois.edu)

[tbarton@uchicago.edu](mailto:tbarton@uchicago.edu)

[marpierc@iu.edu](mailto:marpierc@iu.edu)