

# Identifying Researchers: Some Formal Considerations

Nick Nicholas  
PILIN ANDS transition project

<http://www.pilin.net.au>

# Summary

- PILIN has not deal with Identity:
  - our referents don't talk back
- PILIN thoughts on identifiers align with the use cases presented
- Persistence is about policy and is domain specific
- Interoperability of identifier services
- No such thing as a universal identifier
- Verification of identity and identifier

# PILIN vs. Identity

- PILIN *modelling* applies to any identifiers
  - PILIN thinking can inform identifying researchers
  - PILIN workflows did not account for identity
- PILIN *development* applies to identifiers of intellectual output
  - IDs for intellectual output essential to claims of identity (" author of...")
  - PILIN solutions relevant to research identity
- Mobility use case for researcher ID a known PILIN pain point

# Caveat: Identifier $\neq$ identity

- Identifier is Name associated with Thing; association is captured as resolution data
- Identity are claims about the researcher.
  - Identity is only mediated by the identifier.
- Identifiers can identify abstractions
  - resolving to representations of the abstraction
  - E.g. *Hamlet* vs. PDF of one version of *Hamlet*
- Identity is one representation of the abstraction “researcher”
- Researcher can control the representation.
  - They can’t control the identifier itself

# Researcher ID requirements

- Researcher identifiers should be (*xref to position paper reqts & use cases*):
  - **Persistent** (7)
  - Easily Citable (8)
  - **Interoperable** (6, 9, UC 6)
  - **Verifiable** (4)
  - **Globally registered [Universal?]** (10, UC3, UC6)
  - Leveraged for info services (11)
  - **Unique** (12)

- Relations between researcher and intellectual outputs should be:
  - **Verifiable** (5)
  - Exposed (12, UC 4)
  - **Authoritative** (3)
  - **With Multiple Authorities** (UC 5)
  - **Authorised by researcher** (1, 2)
  - **Universal scope** (UC 1)
  - Aggregated (UC 2)

- Not necessarily for 1000 years
  - We're having trouble enough with 6 months!
- Persistence = available for a defined time period
  - Defined by context
  - Exposed as policy guarantee (trustworthiness)
- Persistence of the identifier is not just the researcher's responsibility!
  - Role of identifier provider

<http://www.pilin.net.au>

# Opaque?

- Opaqueness copes with change in referent
  - to assure persistence
- Opaqueness copes with ambiguity
  - to assure uniqueness
- Scope of uniqueness and persistence determined by context
- Risk mitigation: People less at risk of name change, but impact is greater
- Name + disambiguators *may* be acceptable label
- But not e-mail: institution or ISP-specific.
  - Greater risk of change over time.

<http://www.pilin.net.au>



# Global (Universal?)

- There are Fewer Identifiers (OpenID, SSO)
- There is no One Universal Identifier
  - Anyone can make an identifier for you
  - There may be One Preferred Identifier for a system; But other identifiers don't go away
- Researcher identity management reconciles multiple identifiers for same researcher
  - each with its own claims about identity
  - So no Universal Scope of identifier either
- Solutions
  - Global (?) registry of equivalent IDs
  - Discovery of equivalence:  
Comparable association data (name+)

<http://www.pilin.net.au>

- Identifiers are commonly used as parameters to services
  - And often confused with the services
- Identifiers (strings) are interoperable
- The services may not be
  - One Transport (HTTP)
  - But not one interface or info model

# Verifiable (Claims)

- Not only identifier manager gets to make claims about researcher identity
  - The researcher can too (user-centered identity)
  - And so can anyone else.
- Researcher can decline from lending their authority to claims
  - Claims must establish authority ("who says so?")
  - Claims should be verifiable
- Authority is a means to the end of verification
  - We only care who says so, if the "who" is reliable
  - The researcher should be reliable about themselves
  - We may verify the claim independently anyway (audit)

<http://www.pilin.net.au>

# Verifiable (Identifier)

- Verifying referent not same as verifying claims about referent!
  - Are these the publications of Dr Cornelius Zaius?
  - Is `openid.aol.com/dr_zaius` = Dr Cornelius Zaius?
- Identity doesn't work without verified identifiers
- The mechanisms for verification are different for people (authentication) and PDFs

<http://www.pilin.net.au>

- Persistence is about policy
  - and is domain specific
- Interoperability of identifier services
  - not identifiers
- No such thing as a universal identifier
  - Register or Discover multiple IDs & relations
- Verification of identity and identifier
  - Distinct activities