CrossRef
Contributor ID

STM Innovations
London
December 4, 2009
#stmi09

Geoffrey Bilder
Director of Strategic Initiatives
CrossRef
Contributor ID

STM Innovations
London
December 4, 2009
#stmi09
Author identity: The Shape of the Problem

STM Innovations
London
December 4, 2009
#stmi09
Why do we care?
Trust

...belief in the veracity, good faith, and honesty of another party, with respect to a transaction that involves some risk.

Trust

...belief in the veracity, good faith, and honesty of another party, with respect to a transaction that involves some risk.

Provenance

the place of origin or earliest known history of something: an orange rug of Iranian provenance.

- the beginning of something's existence; something's origin: they try to understand the whole universe, its provenance and fate.

See note at origin.

- a record of ownership of a work of art or an antique, used as a guide to authenticity or quality: the manuscript has a distinguished provenance.
Brand

Is it relevant?
Is it good?
Is it important?
Brand
Is it relevant?
Is it good?
Is it important?

provenance infrastructure
provenance infrastructure

* What is it?
* Where can I get it?
* What refers to it?
* What does it refer to?

* What has been done to it?
* Who provides stewardship of it?

* Who do we credit it to?
* What are their credentials?
Given a work, tell me who is responsible for it and describe the nature of that responsibility.
Programmer
Programmer

Compiler

Tuesday, 8 December 2009
Given a name, tell be what works they have contributed to and describe the nature of the contributions.
Given a name, tell me what other names are “related” to the first name and tell me the nature of that relationship.
Stakeholders?

- Researchers
- Libraries
- Institutions
- Funding agencies
- Intermediaries
- Publishers
So what are the problems?
“Security & Access Control”
vs.
“Knowledge Discovery”
Wei Wang
Terminology
Subject
Identifier
Geoffrey Bilder

Identifier
Geoffrey Bilder
SS# 447766410380

Identifier
Geoffrey Bilder

SS# 447766410380

Identifier
Geoffrey Bilder
SS# 447766410380

gbilder
gwbilder
geoffrey.bilder
dgoffrey_bilder
dgoff

ggeoff

ggeoff23

Identifier
Geoffrey Bilder

SS# 447766410380

gbilder
gwbilder
geoffrey.bilder
geoffrey_bilder
geoff
geoff23

Identifier

gbilder.com
Profile
doi: 10.1007/BF01829975

Score: 100
doi: 10.3998/3336451.0009.101
Persona
Persona
Persona
Credential
Protocol
X.500 Protocol
One more term...
Assertion
Geoffrey Bilder is Creditworthy
Geoffrey Bilder is a US Citizen
Geoffrey Bilder is a UK Citizen
Geoffrey Bilder is authorized to drive a car in the UK
Geoffrey Bilder has bought 1,151 songs from iTunes
I play the electric ukulele
I “follow” Allen Renear on Twitter
I like Sichuan food
I am the author of article X

Assertion
Externally asserted identity
CREDIT HISTORY
Self-asserted identity
Paradox

Origin ofAssertions

Externally-asserted
Exponentially-asserted

Subject’s Trust
External Trust

Trust

Self-Asserted

Tuesday, 8 December 2009
Hybrid-Asserted Identity
Automated suggestions

Proxy Claims:

Leigh claims Geoffrey wrote this article

Self Claims:

Geoffrey claims Geoffrey wrote this article

Verified Claims:

Geoffrey claims Geoffrey wrote this article, and the "Journal of Psychoceramics" confirms that claim

Verified Proxy Claims:

Geoffrey (who has already been verified as an author of this article) claims that Amy was also a contributor to this article
Scope
International Interinstitutional Interdisciplinary
4 Centuries of Citation

- The oldest content registered in CrossRef thus far dates from 1665...Philosophical Transactions
- Chain reaction effect in retro-digitization is driven by linkability of older content
4 Centuries of Citation

- The oldest content registered in CrossRef thus far dates from 1665...Philosophical Transactions
- Chain reaction effect in retro-digitization is driven by linkability of older content

<table>
<thead>
<tr>
<th>Century</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1600s</td>
<td>1,674</td>
</tr>
<tr>
<td>1700s</td>
<td>4,436</td>
</tr>
<tr>
<td>1800s</td>
<td>381,187</td>
</tr>
<tr>
<td>1900s</td>
<td>15,750,357</td>
</tr>
<tr>
<td>21st C</td>
<td>8,031,503</td>
</tr>
</tbody>
</table>
gbilder.com
Scale
200K DOIs Per Month
~5 Authors Per DOI
96-97%
6 Million Researchers
6.7 Million Graduate Students
Finally...
Principles for the design of a contributor identifier...
Principles for the design of a contributor identifier...

- It should be designed to support the creation of a clear and unambiguous record of scholarly communication.
Principles for the design of a contributor identifier...

- It should be designed to support the creation of a clear and unambiguous record of scholarly communication.

- It should transcend discipline, geographic/national and institutional, boundaries.
Principles for the design of a contributor identifier...

- It should be designed to support the creation of a clear and unambiguous record of scholarly communication.

- It should transcend discipline, geographic/national and institutional, boundaries.

- It should be designed to identify “contributors”, not just "authors".
Principles for the design of a contributor identifier...

• It should be designed to support the creation of a clear and unambiguous record of scholarly communication.

• It should transcend discipline, geographic/national and institutional, boundaries.

• It should be designed to identify “contributors”, not just "authors".

• It should support reliable attribution in both formally and informally published literature.
Principles for the design of a contributor identifier...

- It should be designed to support the creation of a clear and unambiguous record of scholarly communication.

- It should transcend discipline, geographic/national and institutional boundaries.

- It should be designed to identify “contributors”, not just "authors".

- It should support reliable attribution in both formally and informally published literature.

- It should be “open” whilst complying with the privacy requirements of the individual as well as of various legal jurisdictions
Principles for the design of a contributor identifier...

• It should be designed to support the creation of a clear and unambiguous record of scholarly communication.

• It should transcend discipline, geographic/national and institutional, boundaries.

• It should be designed to identify “contributors”, not just "authors".

• It should support reliable attribution in both formally and informally published literature.

• It should be “open” whilst complying with the privacy requirements of the individual as well as of various legal jurisdictions

• It should be persistent. This is both a “technological” imperative and a “social” imperative.
Principles for the design of a contributor identifier...

- It should be designed to support the creation of a clear and unambiguous record of scholarly communication.
- It should transcend discipline, geographic/national and institutional, boundaries.
- It should be designed to identify “contributors”, not just "authors".
- It should support reliable attribution in both formally and informally published literature.
- It should be “open” whilst complying with the privacy requirements of the individual as well as of various legal jurisdictions.
- It should be persistent. This is both a “technological” imperative and a “social” imperative.
- It should be controlled by the contributor.
Thank You

gbilder@crossref.org