Using identifiers in patent databases

Osmat A. Jefferson
orcid.org/0000-0002-0607-8728

Joint initiative between Cambia and QUT
Innovation cartography

What?
*Ex-ante* mapping of **knowledge artefacts** to mitigate risks, bring clarity into this process, and guide decision-making

How?
Contextual and dynamic mapping
One to many relationships versus one to one relationships

Why?
Open instead of closed to build confident and trusted partnerships and collaborations
Current silos of Knowledge

Scholarly publications
66,255,368
10%
Open access

Patent families
55,671,970
100%
Open access

Owners of patent rights

Manufacturing standards

Regulatory information

Databases
Linking and establishing contextuality

- Scholarly publications: 66,255,368 (10% Open access)
- Patent families: 55,671,970 (100% Open access)
- Owners of patent rights
- Manufacturing standards
- Regulatory information
- Databases
Ambiguity in the name, work, institution, and metadata

>600 variations


Kohler; Milstein: 'Continuous cultures of fused cells secreting antibody of predefined specificity' NATURE vol. 256, 1975, pages 495 - 497

KBHLER ET AL., NATURE vol. 256, 1975, page 495

KBHLER; MILSTEIN NATURE vol. 256, 1975, page 495

KOENIG ET AL., NATURE, vol. 256, 1975, pages 495 - 497


KOENIG; MILSTEIN NATURE Bd. 256, 1975, Seiten 495 - 497

KOHNER ET AL., NATURE, vol. 256, 1975, pages 495 - 497

KOINER ET AL. NATURE vol. 256, 1975, pages 495 - 497

KOEHLER ET AL., NATURE, vol. 256, 1975, pages 495 - 497


Lens.org

Using **open persistent identifiers**, the Lens links patent knowledge artefacts (LensID), joins scholarly literature with patent literature to show where scholarship has influenced industry (DOIs and PMIDs), and enables real time exploration of their interactions at institutional level and individual’s work level (ORCID ID).

Lens.org hosts 100M+ patent records from ~95 jurisdictions,

- 35 M non patent literature records from more than 30 patent offices
- > 271 M patent sequences from 17 jurisdictions integrated with other genomic resources, including the OMIM database

(PatSeq is the world’s first comprehensive public database)
Open persistent identifiers

ORCID ID
orcid.org/0000-0002-0607-8728

LENS.ORG
lens.org/029-768-644-186-340

PubMed
PMID is a unique number assigned to each PubMed record.

27 Million citations

PubMed Central
PMC is a free digital archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health (NIH)

Crossref

Global Legal Entity Identifier Foundation
LensID

Patent documents are generally not copyrighted. They are present on several websites in various degrees of utility. Provided in different formats and identified with different codes. To capture the constellation of related knowledge (Knowledge artefacts) of a patent record, we created a LensID, an open persistent identifier.

LensID now includes - but not limited to- bibliographic information, filing history, inventor attested ID (if available), citations, associated sequence listing files, and legal status.
LensID_15 digit congruent with ORCID and ISNI schemes

Prefix. The hyperlink should point to https://lens.org/029-678-644-186-340 and this applied to URLs in web service data as well. Using HTTPS allows the user to *privately* request documents from lens.org

Code to generate LensID is posted at https://www.lens.org/support/help-resources/basic-help/the-lens-id/
Linking ORCID works to patent citations maps influence of an individual’s work on industry.
PatCite- Joining scholarly work (across many levels) to patent citations to map influence on industry and innovation

We are sure that science and technology influences and enables industry. But how do we map this and measure it? And how do we improve it?

The Lens has mined the global patent literature for citations of scholarly published work, and provides free and open tools to explore which articles influence what patents, held by what institutions. Investors and funding agencies can find whether the science they fund enables outcomes for society through new inventions and products.

Using this toolkit you can track, filter, sort, and link scholarly articles that are cited in patents, examine the citations, and discover new partners and collaborators. Influence Mapping provides an evidence base to guide decision making and enables improved public policies and institutional practices.

PatCite is Coming Soon.
Extending LensID to innovation metadata and KAs

Über Datensätze
for Patents, Scholarship, Legal Entities, Humans

A master record for each entity in which we can aggregate the link outs and related information about a particular knowledge artefact (KA). Here, we propose adding letter tags in the prefixes of the ID to indicate whether the ID is patent related, human, scholarly, or legal entity related;

H- LensIDs refer to every listed inventor, every listed author
P- LensIDs refer to every patent record we have
S- LensIDs refer to every scholarly record (DOIs, PMIDs or others).
L- LensIDs refer to every Applicant and Institution and Owner.

Prefixes are NOT part of the ID but simply to enable users to cluster and group IDs. The intention here is to build an open mapping of the IDs themselves and a logging system that allows grouping and disambiguation that is completely transparent and freely reversible.
Example of open action-history logging of identifiers

New ID is used to make a new group
This open mapping is transparent and maintains a logging history
Acknowledgments

- Funding from Bill & Melinda Gates Foundation, Gordon and Betty Moore Foundation, Lemelson Foundation, NCBI/NIH, USPTO, Qualcomm, NICTA, and Syngenta
- Arun Sharma (financial support from QUT)
- Patent offices, University of Cambridge, and others