

Linking Patents to Indicators

and Patent Metrics

Jeffrey D. Oldham

2016-05-20

Link patent metrics to innovation indicators

Patent metrics summarize sets of patents.

- scopes, family members, backward and forward citations, dates, ...
- Lower barrier to patent research.

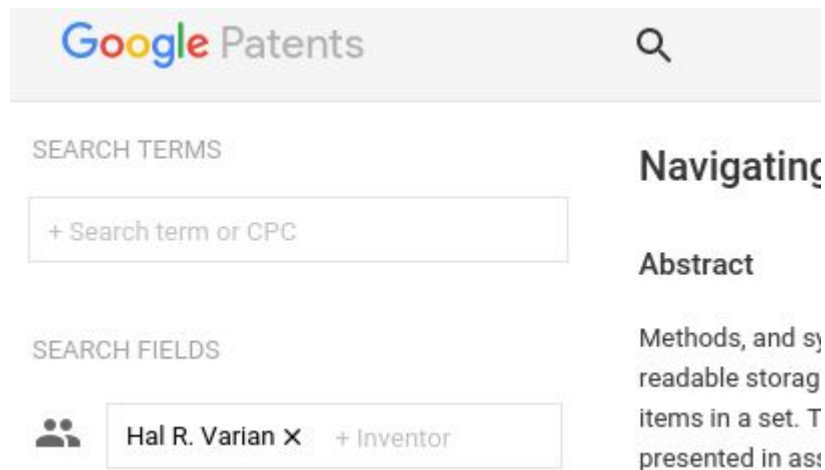
Link free metric database to other indicators.

- Researchers contribute data sets.
- Explore connections among indicators.

Accessing patents

Currently: individual patents

[Google Patents](#), [USPTO PatentsView](#)



The image shows a portion of the Google Patents website. At the top, the 'Google Patents' logo is on the left and a magnifying glass search icon is on the right. Below the logo, the text 'SEARCH TERMS' is followed by a text input box containing the placeholder '+ Search term or CPC'. Further down, the text 'SEARCH FIELDS' is followed by a section for filters. One filter is active, showing a person icon, the text 'Hal R. Varian X', and a '+ Inventor' button. To the right of the search area, the word 'Navigating' is partially visible, followed by the word 'Abstract' and a snippet of text: 'Methods, and sy readable storag items in a set. T presented in as:'.

+ Metrics

Sets of patent summaries

US20140019868A1:

publication date: 2014-01-16

scope: G06F, G06F

claims: 33

backward citations: US20060059440A1, ...

forward citations: US20140229869A1, ...

family members: EP2873053A2, ...

Ex: Time-order G06F scope patents.

Metrics

publication number: CN102673696A,
EP2032523B1, US3540328A, DE429373C,
WO2015042844A1

is minor publication: true, false

publication kind: application, grant

scope: H01L, A61J, B01F, G06F, Y10T

family members: WO1996006179A1,
US9140531B2

number of claims: 0, 20, 56

backward citations: NULL, US5797176A,
JPH0368149A

forward citations: USD747137S1,
CN102669786A

publication date: 1991-02-21

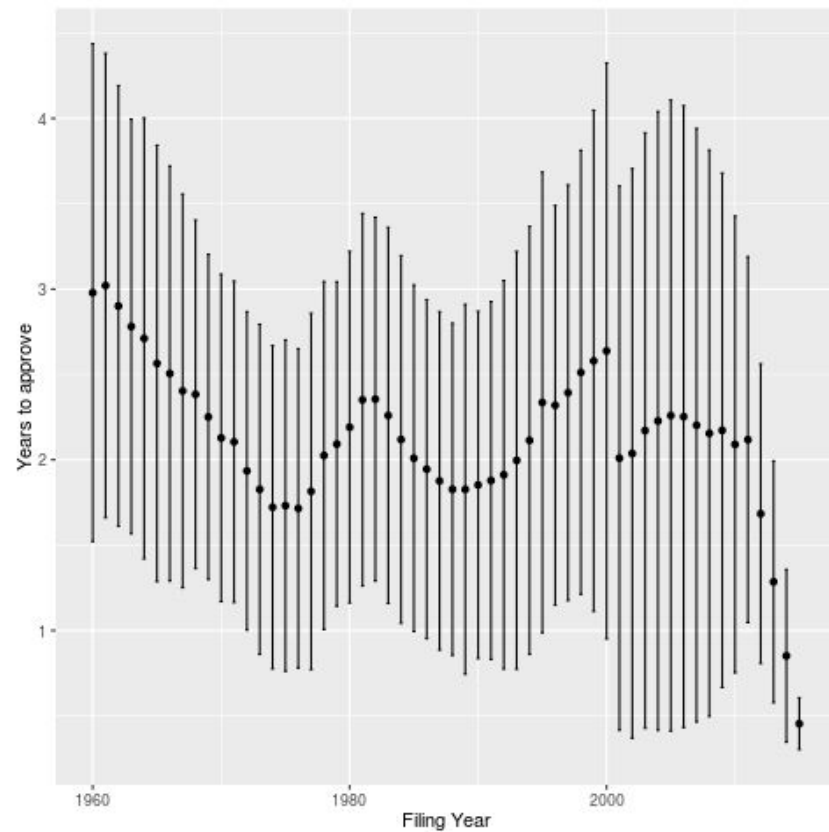
filing date: 1989-09-08

priority date: 1989-08-13

Herfindal (1 - generality): 1.0, 0.7105, 0.441

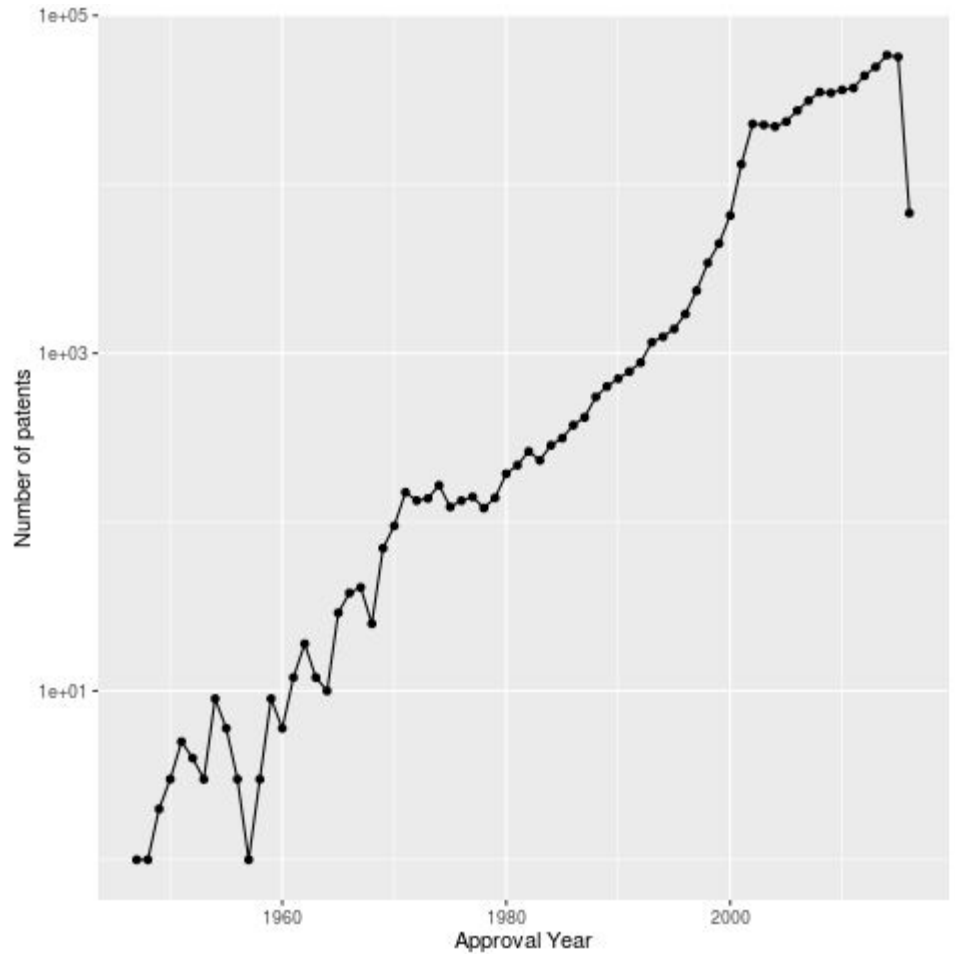
(Precise metric definitions will be published.
(**purple** denotes a repeated field.)

Are US patent approvals slowing?



Is scope G06Q growing?

*data processing systems or
methods*



Metrics ease patent cleaning

Ex: publication date - filing date > 50 years:

US4702526A	1987-10-27	1885-07-29	wrong USPTO filing date
US38739A	1963-06-02	1862-12-27	Google OCR bug
USPP3117P	1972-04-18	1879-03-09	Google OCR bug

"In order to gain meaningful understanding ..., it is necessary in many cases to link data from different sources."

[Hall+Jaffe, 2012, p.38]

Link metrics to other patent data

- + normalized inventor, assignee data [[Fleming](#), McCue]
- + normalized geos
- + [PAIR](#) blocking data [Fleming]
- + patent case data ([PTAB](#), [PACER](#))
- + corporate hierarchies

Link metrics to non-patent data

- + census data — patents' impact on quality of life
- + inventor–school(s) — training effectiveness
- + inventor–school funding source — NSF STAR Metrics
- + gov. R&D funding — program effectiveness
- + company sales data — financial impact of design patents [[B-M+T 2011](#)]
- + sector sales data — impact of innovation on sales
- + inventor citizenship — effect of immigration policy

Implementation sketch

Patent metrics available via Google BigQuery.

- SQL via interactive WWW or [programming language](#).
- [\\$0.00](#) for 1st TB of data analyzed.

Link other data via patent number, inventor ID.

Store in [Google Cloud Storage](#).

- [\\$0.02 / GB / month](#)

Implementation does not yet exist.

Linking Patents to Indicators

and Patent Metrics

Jeffrey D. Oldham

2016-05-20

Patent metrics CSV data size

metrics (<i>non-repeating</i>)	0.8GB (4.6GB uncompressed)	71M patents
family members	1.8GB (6.1GB)	253M members
scopes	1.3GB (4.0GB)	235M scopes
backward citations	1.6GB (5.2GB)	215M citations
forward citations	1.6GB (5.2GB)	214M citations
total	7.1GB	

PATSTAT

[Online patent database](#) with SQL.

Advantages:

- run by European Patent Office
- clean data
- [documentation](#) and [sample queries](#)

Disadvantages:

- [cost](#)
- extensible?