



BRETTTERBLOG

ÜBER UNS

WARUM BRETTTERBLOG?

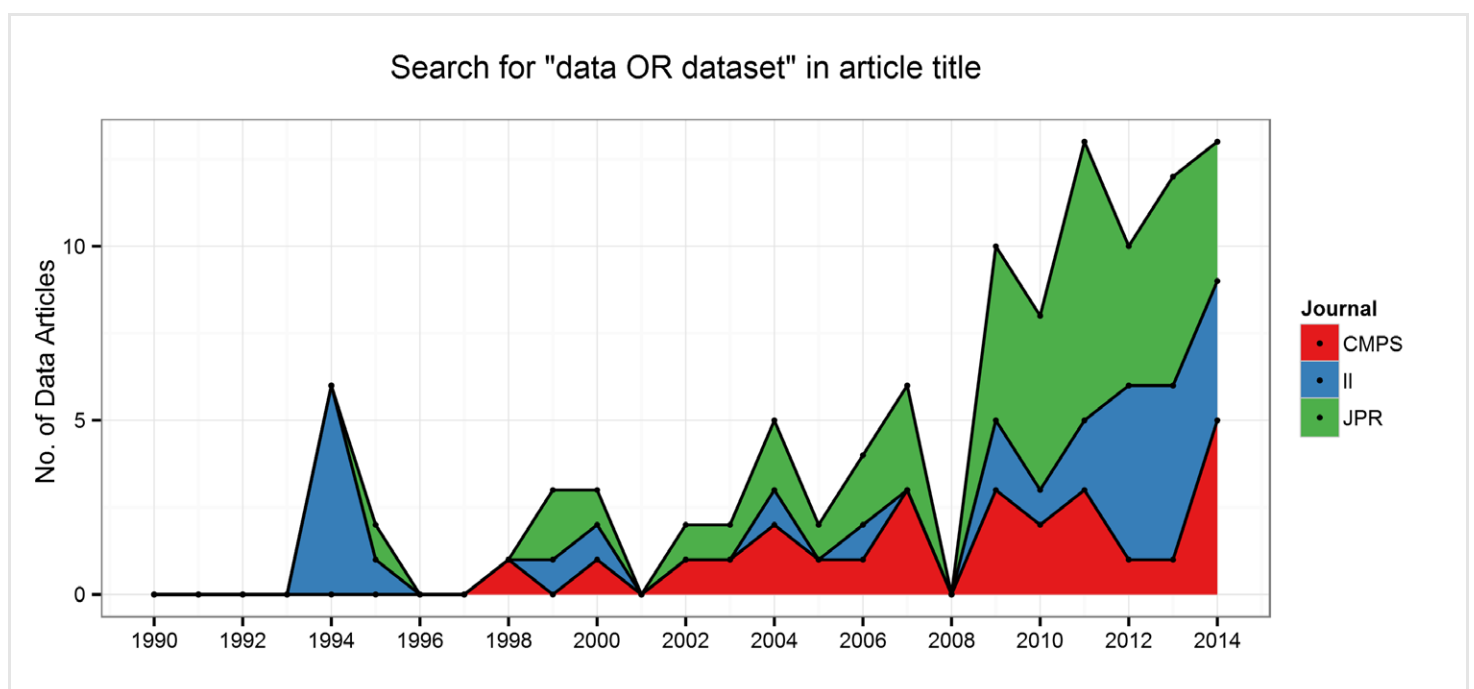
Perspektiven auf globale Politik

ARE WE DROWNING IN CONFLICT DATA?

8. April 2015 · von Felix Haass · in *Debatte* · [Hinterlasse einen Kommentar](#)

I'm probably not alone in observing that there seems to be an increasing number of data articles being published in the field of conflict studies and IR. Together with some colleagues, I'm even preparing one myself at the moment! Is that perceived increase in data publication actually measurable? And does it indeed amount to "drowning"?

The answer to the first question is certainly yes. Scraping data from Google Scholar, I searched the three major IR journals that regularly publish data feature articles (*Journal of Peace Research*, *Conflict Management and Peace Science*, and *International Interactions*) for the terms "data" and "dataset" in their article titles since 1990. I then aggregated the counts by journal name and year. This reveals the following trend:

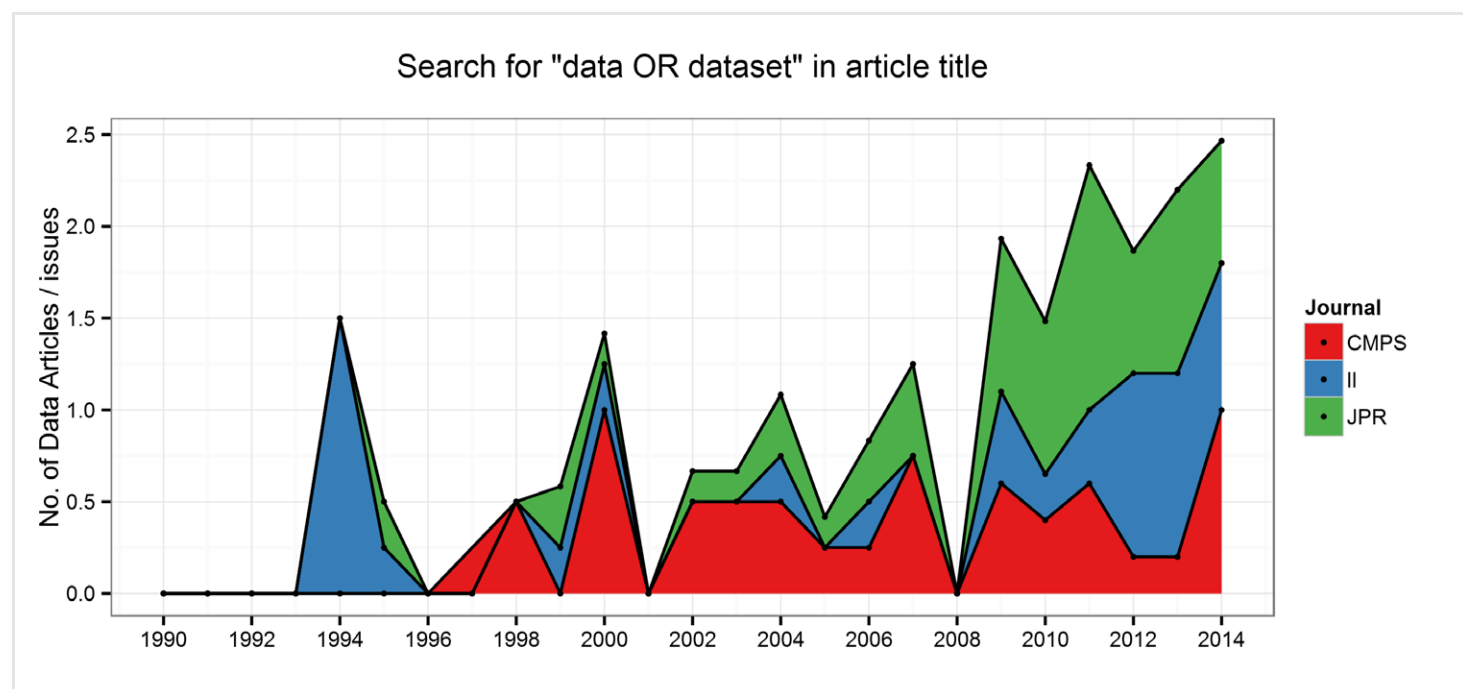


Now, this is only a quick and dirty search which is certainly flawed. For one, it is likely that my search counts articles as "data

publications” when they are only re-examining existing data. Plus, my search does not include articles who do present new data, but do not include the term “data” or “dataset” in their article titles. *JPR*, for instance, publishes an annual update to its **conflict data**. My data collection records no published dataset in 2008, however, but a quick manual search reveals that UCDP did publish their update in 2008. But they named it “**Dyadic Dimensions of Armed Conflict, 1945-2007**” which excludes it from my search criteria.

Also, journals may have increased their publishing rate. If they publish now 6 issues per year whereas they used to publish only 4 back in the 90s, the chances for data getting published increase. Finally, and probably most importantly, not every new dataset is published in a separate data feature article. I’d say most of new datasets are introduced in original research articles—but it is difficult to capture those programmatically.

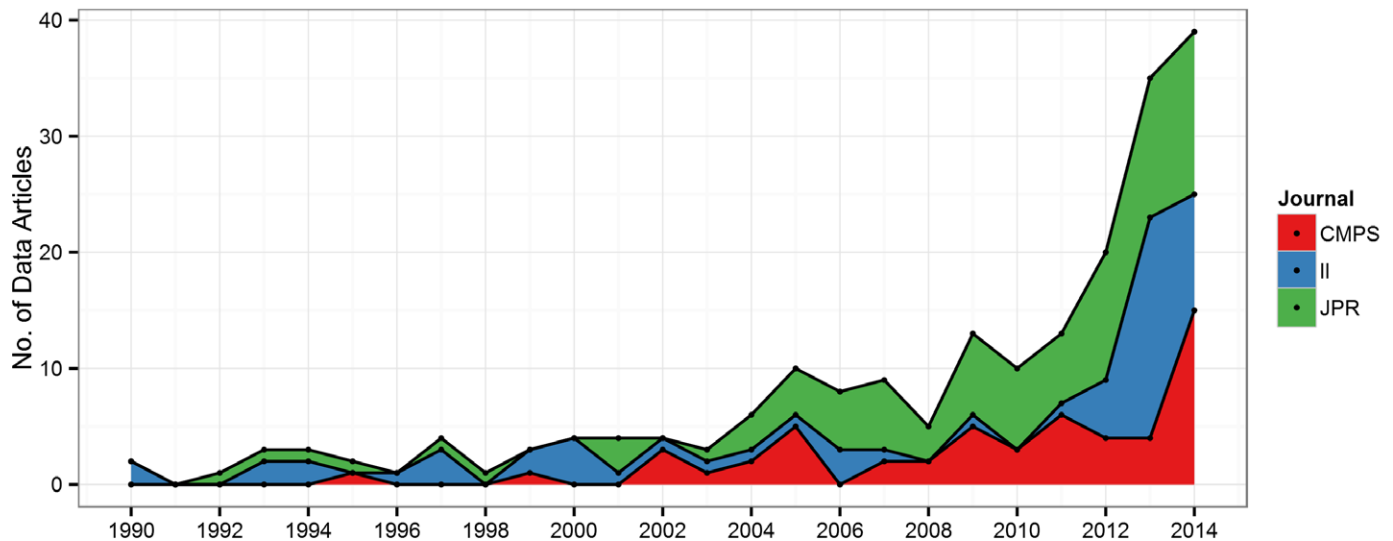
To somewhat correct for the fact that increased data publication may simply reflect an increase in overall journal output, I coded the number of annual issues per volume and “normalized” the number of data articles per number of annual issues. This procedure reveals a somewhat similar trend—although the earlier years get a little bumped up if we account for the lower number of issues published in the 90s and early 00s:



The main reason is that all journals increased their number of issues per year over the last 20 years. The increase in data articles was accompanied by an increase in journal output which then results in the same pattern as simply counting data articles per year.

I also at least wanted to try to somehow capture those articles that introduce original datasets in the context of a regular research articles. I therefore searched for the term “new data” in all articles (body + title) to get an idea of how frequently authors referred to somehow newly collected data.

Search for "new data" in article title & body



This reveals a much steeper increase than the “mere” data feature articles, especially in the last four years, but I’d assume the number of false positives is also much higher. A “normalized” plot that accounts for varying issues/year looks similar, again which is why I won’t reproduce it here. It suffices to say that the “data revolution” in IR and conflict is reflected in the number of articles that publish data or that refer to “new data”. It would probably be even more steep if we included articles from other Journals, such as *Journal of Conflict Resolution*.

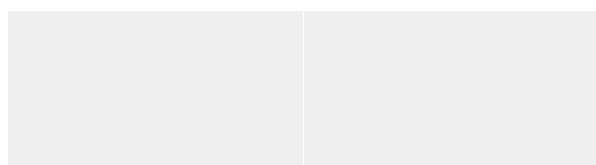
But does that mean we’re “drowning” in conflict data? Answering that question is a much more difficult task than simply scraping the raw numbers from Google Scholar. It is so difficult because it entails a whole string of follow-up questions: is there such a thing as “too much” data? Who is going to analyze all of these data? And, with so much data being produced, can we ensure that the data are of high quality (see, for instance, [this JPR piece](#) on data quality)? Plus, growth rate of data articles looks almost exponential—but from all what we know about exponential growth, should we expect a limiting factor to kick in at some point? Is the increase in information reflected in increase in actual *knowledge* about peace and conflict?

I obviously can’t answer all of these questions in this quick blog post. The growing availability of data is certainly a good thing in that it enables us to understand much more about conflict from a social science perspective (and probably from a policy perspective, too). At the same, this data explosion should also make us pause a moment and think about where we are going with this, what we can and cannot do with this amount of data, and—maybe most importantly—what we *should* and *shouldn’t* do with this flood of information.

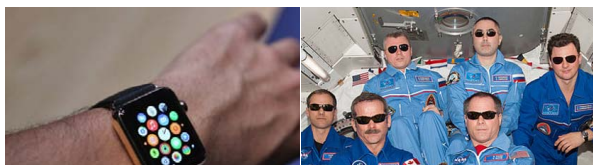
Postscript:

I’ve uploaded the R script to download the data from Google Scholar and reproduce the plots [here](#). Feel free to tinker with it yourself. I’m grateful to [Kay Cichini](#) for writing the excellent [GScholarScraper](#) function in R which I modified a bit to allow for searching within specific Journals.

Crossposted from the [author’s website](#).



Apple Watch has 8GB of NASA wants astronauts



storage, but you can't use

to wear smart glasses



10 Most Luxurious Hotels in the World



The 13 Worst Foods to Eat at Night

by Gravity

Bewerten:

★★★★★ ⓘ 1 Vote

Teilen mit:

Twitter Facebook E-Mail Google

Lade...

Ähnliche Beiträge

IB Online (1/10): Eine kleine Netzschau
In "Links"

Anstelle einer Netzschau: Links aus dem Sommerloch
In "Links"

IB Online (3/2): Eine kleine Netzschau
In "Links"

Tags: conflict, data, international relations, journals, political science, publishing, R, trend

Kommentar verfassen

Gib hier Deinen Kommentar ein ...

Chinas Entwicklungsbank AIIB: Der Weltbank hässliches Schwesterlein?



Folgt uns per E-Mail:

Schließe dich 893 Followern an

SUCHE

LETZTE KOMMENTARE

- Kommentar zu Strafe als Chance: die Russland-Sanktionen von Ivan Samolovov 23. März 2015
- Kommentar zu Strafe als Chance: die Russland-Sanktionen von Jorge Garzon 18. März 2015

TWITTER

- Are We Drowning in Conflict Data? wp.me/p2tUkv-1b6 | 21 hours ago
- RT @EAProspero: Wenn wir schon dabei sind. @bretterblog, persönlicher Favorit, wenn es um Blogs von dt. Politikwissenschaftlern geht <https://...> | 1 week ago

BLOGROLL

- A View From the Cave
- Aidnography
- ArmsControlWonk
- Chris Blattman
- Conflict & Security
- Göttinger Institut für Demokratieforschung
- Inner City Press
- International Law Observer
- International Security Discipulus
- IR Blog (Berlin)
- ISN Blog ETH Zürich
- Jihadica
- Junge UN Forschung
- Junge Wissenschaft im öffentlichen Recht
- Justice in Conflict

Kings of War

- menschenhandel heute
- Mobilizing Ideas
- Political Violence @ a Glance
- Sakranut
- sicherheitspolitik-blog.de
- Stephen M. Walt
- The Duck of Minerva
- The Monkey Cage
- theorieblog.de
- Theory Talks
- Thought Economics
- UN Dispatch
- University of Notre Dame – Peace Policy
- Völkerrechtsblog
- Why Nations Fail
- World Peace Foundation – Reinventing Peace
- zoon politikon

META

- [Registrieren](#)
- [Anmelden](#)
- [Beitragsfeed \(RSS\)](#)
- [Kommentare als RSS](#)
- [Bloggen auf WordPress.com.](#)

[Bloggen auf WordPress.com.](#)

The Origin Theme.