

Of Festschriften and Preprints

Exploring Understudied Academic Genres

Hanjo Hamann

A. Background and Motivation

How better to celebrate a pioneer of quantitative legal studies than by quantitatively studying his own legacy? To be sure, the following exploration (like all exploratory statistics) can only paint one view of the cathedral that is the Engel legacy. Yet, this particular view may be of general interest beyond the professional network whose members assemble in this festschrift.

But first, a few thoughts about the festschrift as such.

I. Of Cemeteries and Promenades

Festschriften are a German thing. There is no proper English word for them, so the English language in 1898 simply ran with the German term.¹ Even one century later though, not all German law professors were fond of *festschriften*. A German émigré with some outside perspective called them “mass burials of erudition” and “cemeteries of research”,² and a lawyer-and-economist emeritus beheld *festschriften* “as graveyards on which original ideas, coherent thoughts, and masterpieces of style are laid to rest”.³

Coincidentally it was in just the (few) years between these quotes that a new-fangled research unit on law-and-economics was established in the

1 See www.etymonline.com/word/festschrift.

2 *Stiefel*, JZ 1995, 613: “Der Festschriftler [...] hört die Kassandrarufo: Massengrab der Gelehrsamkeit. Er merkt, sein Beitrag ist im Leichenwagen auf dem Wege zum Friedhof der Wissenschaft.”

3 *Wolfram Müller-Freienfels*, cited by *Löwisch* (JZ 1998, 946) and *Münch* (NJW 2000, 3253, 3256 footnote 28): “Festschriften sind wie Friedhöfe, auf denen originelle Ideen, schlüssige Gedanken und stilistische Meisterwerke ruhen.”

former German capital (*Projektgruppe Recht der Gemeinschaftsgüter*).⁴ Its director did entomb quite a few coherent thoughts in the fertile soil of *festschrifts*.⁵

Yet, even before I became his student – and long before I lowered my first coffin of ideas into a *festschrift* – I had already learned to love unearthing *festschrift* articles. Not only do they serve more general interests than most law journals: They also make room for (and almost demand) personal anecdote, creative style, and – quite simply – academic gossip. I gladly comply with tradition by releasing a personal anecdote into the realm of gossip: Throughout my years of research, whenever I requested a library to digitize some *festschrift* article, I invariably perused the volume's entire table of contents and exhausted the 10% copyright limit on digital library reproduction.⁶ There was always something else that piqued my interest. I thus perceive *festschrifts* as a strip mall of sorts, or a French shopping promenade: A quick succession of enticingly dressed windows into academia's haute-couture. Very different styles, some even artsy, but all quite intriguing.

Whence these diametrically opposed views of the *festschrift*? Why did scholars in the 1990s consider them morbid necropolises, while I see them as buzzing bazars? I can only guess that my particular position is that of a digital almost-native at a time when the table of contents of virtually any German book is easily available through the National Library (DNB) catalogue,⁷ while free digitization-on-demand services allow me to obtain a digital copy of any book chapter within a working day. Even though most *festschrifts* are still not available online (unlike this one), or even

4 On the institute's origin, see *Kazemi/Henning*, *Chronik der Kaiser-Wilhelm-/Max-Planck-Gesellschaft zur Förderung der Wissenschaften*. 1911–2011, 2011, p. 700.

5 *Christoph Engel* contributed at least ten such celebratory contributions: For Dietrich Rothoefl (1994), Ernst-Joachim Mestmäcker (1996 and 2003), Klaus J. Hopt (2010), Wernhard Möschel (2011), Martin Bullinger (2011), Paul Kirchhof (2013), Martin Morlok (2019), Christine Windbichler (2021), and Michael Faure (2025). The first two *festschrifts* he co-edited himself.

6 Sec. 60e (5) of the German Copyright Act (UrhG): “In response to individual orders, libraries may, for non-commercial purposes, transmit reproductions of up to 10 per cent of a published work to users [...].”

7 Or through dedicated DNB staff, who digitize missing TOCs for free. Most notably, *Kristina Knull-Schlomann* has been my unflinching pillar to sourcing literature (*festschrift* or otherwise) for the past five years. Let this footnote be a token of gratitude for some 150 emails in which she confirmed the retro-digitization of hundreds of TOCs since our first exchange in October 2020.

sufficiently indexed in article databases, I go article-shopping in them way too often to think of them as burial grounds.

This was all different in the 1990s when the *festschrift-as-cemetery* theorem (as we shall call it) was prominent. It had sprung from two observations: (1) that there were “masses, an inflation of *festschriften*”,⁸ while (2) “legal scholarship and legal practice pass over many a *festschrift* contribution because the *festschrift* is simply not available to them”.⁹ This brings us to this contribution’s topic, which is dear to the hearts of both the author and the dedicatee:¹⁰ Digital access to scholarly writing.

II. Of Electronic Bulletins and Neverprints

The greatest driver for improved access to scholarly articles has, of course, been the Internet. Scientists were among its early adopters when they started sharing drafts electronically in the 1980s – first through e-mail lists, then “electronic bulletin boards”, which preceded our present-day “repositories”. The most influential of these (arXiv) started as the “xxx.lanl.gov physics e-Print archive” in Los Alamos in 1991.¹¹ In the same year, a German economist at the University of Surrey began developing “net-based services for economists” (NetEc),¹² which would become “Research Papers in Economics” (RePEc) just as the aforementioned *Projektgruppe* was established in Bonn. Concurrently with NetEc, financial economists in the US started the “Financial Economics Network” (FEN) in 1993, which would grow into a “Social Science Research Network” (SSRN) the next year. Others followed suit, and now a Wikipedia “list of preprint repositories” lists some eighty databases with millions of arti-

8 Stiefel, JZ 1995, 613: “Der Festschriftler wird gewahrt, daß er Teilnehmer an einer Masse, einer *Inflation* von Festschriften ist.”, citing apocryphal data on “Festschrift-Hypertrophie: 1864–1944: 3 pro Jahr; 1945–1961: 10; 1987: 45; neunziger Jahre um 100.”

9 Löwisch, JZ 1998, 946: “Rechtswissenschaftliche Diskussion und vor allem juristische Praxis gehen vielfach an Festschriftbeiträgen vorbei, [...] weil ihnen die Festschrift einfach nicht zur Verfügung steht.”

10 See, e.g., Engel, Give the Journals Back to the Scientists, JITE 160 (2004), 35–38.

11 See Taubes, Publication by Electronic Mail Takes Physics by Storm, Science 259 (1993), 1246–1248.

12 See Krichel, About NetEc, with special reference to WoPEc, CHEER 11 (1997).

cles.¹³ Institutional repositories have joined their ranks: Some festschrift contributions by *Christoph Engel* were pre-published as drafts in the Publication Repository of the Max Planck Society (MPG PuRe).¹⁴

Even though “e-prints” (as arXiv called them) or “preprints” (as SSRN calls them) are still a quite recent phenomenon – and not widely adopted (or even known) in German legal academia – they are in some sense already outdated. The term “... print” may seem anachronistic at a time when much research really *never* gets printed – or even published.¹⁵ They might better be called “research papers” or “working papers” instead (as RePEc does), as illustrated by the subsequent development of the Bonn *Projektgruppe*: While in 1998 it started publishing its own digital series of “Preprints aus der Max-Planck-Projektgruppe”, this was renamed twenty years later to “Discussion Papers of the Max Planck Institute” – reflecting not just a shift in working language and incorporation status, but also a broader development in the objective of research dissemination.

III. Of Preprint Bibliometrics and the Grey Beyond

Like any evolving social practice, “preprinting” is being studied empirically. Some have found that the practice predates the internet, with the first large-scale preprint platforms in biology existing as early as the 1960s.¹⁶ Today, across academic disciplines, bibliometricians find that “the number of preprints has exponentially increased 63 times in 30 years, although it only accounts for 4 % of research articles.”¹⁷ A comprehensive

13 Available at en.wikipedia.org/wiki/List_of_preprint_repositories. Apart from the three already mentioned, the list contains four newer entries with at least 1 mio. articles: HAL Open Archives (est. 2001), Science Open (est. 2013), OSF Preprints (est. 2017), and Synthical (est. 2023).

14 Of the ten contributions listed in note 5, this applies to at least three of the most recent ones honoring *Hopt* (<https://hdl.handle.net/11858/00-001M-0000-0028-6DA8-C>), *Windbichler* (<https://hdl.handle.net/21.11116/0000-0005-A6A1-9>) and *Faure* (<https://hdl.handle.net/21.11116/0000-000F-AA4B-E>). The repository was only established in May 2009.

15 Research indicates that as few as 41 % of preprints might ever get published, see *Xie/Shen/Wang* (Fn. 17).

16 On the NIH’s “Information Exchange Groups” (1961–1967) see *Cobb*, The prehistory of biology preprints, *PLoS Biol* 15 (2017), e2003995.

17 *Xie/Shen/Wang*, Is preprint the future of science?, *IW3C2 Proceedings* 2021.

review of preprint research is impossible within the confines of this contribution, yet two general study formats seem to emerge.

One explores the *causes* (motivations) for researchers to circulate research ahead of publication: “Preprinting benefits include rapid dissemination of academic work, open access, establishing priority or concurrence, receiving feedback, and facilitating collaborations.”¹⁸ This research has revealed another important catalyst (apart from the Internet) for a recent surge in preprint dissemination: The COVID pandemic.¹⁹

The other strand of research explores the *effects* (reception) of preprints: Do scholars use and cite preprints like published research?²⁰ Is the evidence reported in preprints robust, or does it frequently get revised after peer review?²¹ Which cues do researchers use to assess preprint credibility?²² Can unethical research practices be identified from a linguistic analysis of preprints?²³ (How) Do journalists rely on preprints?²⁴

Despite such studies, preprints remain largely uncharted territory. They are generally considered “graue Literatur”:²⁵ *Grey* literature which is neither *white* (published conventionally and catalogued by libraries) nor

18 Sarabipour et al., On the value of preprints, PLoS Biol. 17 (2019), e3000151.

19 Glymour et al., Counterpoint: Preprints and the Future of Scientific Publishing, Am. J. Epidem. 192 (2023), 1043.

20 Bertin/Atanassova, Preprint citation practice in PLOS, Scientometrics 127 (2022), 6895: “authors make use of different lexical content when citing preprints compared to the rest of citations.”

21 Nelson et al., Robustness of evidence reported in preprints during peer review, Lancet Glob. Health 10 (2022), e1684–87: “the correlation between estimate values before and after review was high (0.99) and there was no systematic trend.”

22 Soderberg/Errington/Nosek, Credibility of Preprints, Royal Soc. Open Sci. 7 (2020), 201520: “cues related to information about open science content and independent verification of author claims were rated as highly important [...] peer views and author information were rated as less important.”

23 Teixeira da Silva, “Tortured phrases” in preprints, Curr. Med. Res. & Opin. 39 (2023), 785: “Whereas some ‘tortured phrases’ might represent simple mistranslations, in other cases, an abundance of such terms in a single preprint might reveal a more serious ethical issue, such as the undeclared use of a paper mill or an unprofessional editing service.”

24 Fleerackers et al., Science in motion, PLoS ONE 17 (2022), e0277769: “journalists approach the decision to cover preprints as a careful calculation [...] journalists also reported barriers to covering preprints, as many felt they lacked the expertise or the time required to fully understand or vet the research.”

25 The term’s origin is unclear, but surely older than the first reference I found: Ullner, Afrika-Dokumentation, Africa Spectrum 4 (1969), 38, 39, 42.

black (publically unavailable and uncatalogued, like corporate R&D);²⁶ there is now a whole ecosystem around it.²⁷

One type of grey literature is institutional literature such as the aforementioned “Discussion Papers of the Max Planck Institute for Research on Collective Goods Bonn”. I am not aware of previous research on this body of 27 years of contiguous research. So, as a contribution both to preprint bibliometrics and to charting the *Engel* legacy, I set out to explore the preprint series of his institute since 1998.

B. Methods and Data Preparation

I. Data Source

I downloaded the metadata of the preprints and discussion paper series from the Max Planck Society’s institutional repository (PuRe) on 28 January 2025. Each paper of the series was recorded with its year and document number (in the format [YYYY]/[NN]), authors, title, keywords, abstracts, and URL of full text pdf. There were no metadata on the papers’ discipline of origin, which to code manually would have been unreliable at best, and epistemologically unjustifiable at worst. So all papers were treated as coming from the same general field of research, i.e., *collective goods research*.

This dataset consisted of 605 repository entries. To prepare them for analysis, data were first transformed into a Stata dataset, and tested for internal consistency. I discussed any inconsistencies with the MPI staff member in charge (*Marc Martin*) whose support I gratefully acknowledge. Several errors in the repository were corrected in this process, and previously unavailable full text documents were reuploaded.

26 See GDR librarian *Schmidmaier*, Ask No Questions and You’ll be Told No Lies, LIBRI 36 (1986), 98, 101: “Grey literature’ is that in the libraries which is not ‘white’ (available, catalogued, classified) or ‘black’ (not available, unknown, not obtainable, secret).”

27 See the Grey Literature Network Service “GreyNet” with its International Journal on Grey Literature simply called “The Grey Journal”, and the “International Conference on Grey Literature” of at least 25 instalments.

II. Completeness and Deduplication

I ensured that data were complete by comparing a list of the paper numbers with an auto-generated list of running numbers for each year. Several gaps were identified, of which only three remained after data preparation: Preprints 2007/23, 2011/27 and 2011/29 were never published, their slot had merely been reserved for one of the institute's directors (*Martin Hellwig*). In addition, preprint 2007/11 had been replaced by two documents: 2007/11a and /11b. Conversely, three preprints were published outside the running count: 2004/1(a), 2007/4b, and 2007/6b.

In addition to ensuring overall completeness, I also tested whether metadata for each entry was complete. Whereas some data (like title and authors) were available for every preprint, others had not been recorded consistently and could not easily be interpolated or reinserted. In order to avoid biases, I did not use incomplete data in subsequent analyses. This included upload date (missing in 84 % of records), keywords (missing in 35 % of records), and abstract (missing in 16 % of records).

I then checked for metadata duplicates. There were none in the strict sense, but in a few instances, the repository entry noted a preprint being “updated by” a different one from the same series. In these cases, the earlier paper was treated as a duplicate, so as to include only the latest version in some analyses. I proceeded likewise for all papers with identical titles (all but one of which also had identical authors).²⁸ Some of these papers acknowledged, in their full text, the earlier version which they superseded, others did not.

Overall, I labelled 13 papers as duplicates.²⁹ Unfortunately there was no easy way to verify whether other papers superseded one another under altered titles, or whether even similar-sounding titles³⁰ indicated duplicates. This might be possible by semantically comparing abstracts with one another, or even by computing a similarity score for pairs of full text documents; both were infeasible for this initial exploration.

28 Paper 2019/9 had one author less than its previous version 2017/18.

29 See footnotes 34, 35 below.

30 Compare, for instance, the title of 2008/29 (“Competition Policy and Sector-Specific Regulation for Network Industries”) with 2018/7 (“Competition Policy and Sector-Specific Regulation in the Financial Sector”), or 2014/9 (“Financial Stability, Monetary Policy, Banking Supervision, and Central Banking”) with 2015/10 (“Financial Stability and Monetary Policy”).

III. Full Text Acquisition

For most repository records, a full text had been uploaded. For others, their metadata included external URLs, albeit inconsistently: Only 82 % of records featured an external URL, which in three quarters of the cases (457) was of the form *http://www.coll.mpg.de/pdf_dat/[year]_[##]online.pdf*. This uniform scheme allowed for easy interpolation, which worked in every instance except five.³¹ For these I found URLs on different domains, except in the earliest case (1999/9), whose author (*Christoph Knill*) kindly provided another document version which he assured me was identical to the one submitted to the present paper series.

After assembling the URL list, I downloaded each document and extracted their page counts, which I used for subsequent analyses. I did not use the preprints' contents yet due to the difficulty of extracting clean contents from pdf files with vastly different layouts. While they might be used for a variety of semantic comparisons and interesting measures (discipline-specific vocabulary, number of references, existence or absence of economic theory, statistical tests, etc.), this was beyond the scope of an initial exploration.

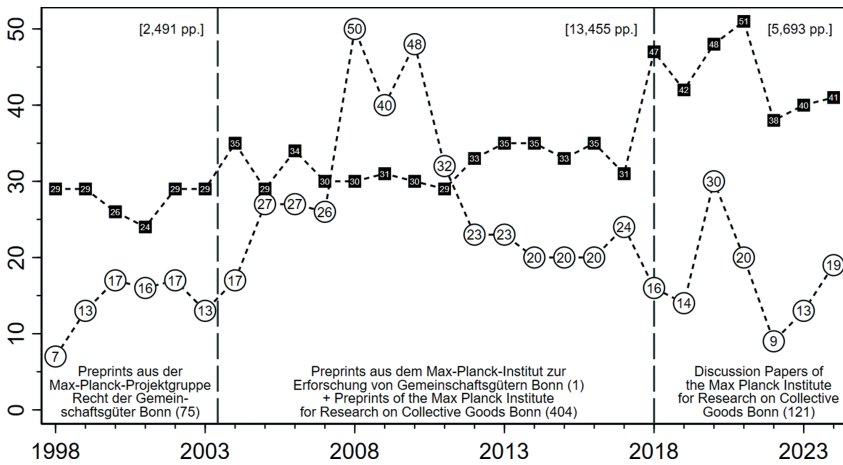
C. Data Exploration

I. Distribution: 21,639 pages over 26 years

As a first exploratory analysis, I show the number and length of papers over time (including duplicates) in graph 1 below.

Vertical lines divide the graph into three phases: In the first (up to 2003/5), 75 documents of just under 2,500 pages had been called "Preprints aus der Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter Bonn". Then came a single document (2003/6) under the series title "Preprints aus dem Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern Bonn" before the series was again renamed to "Preprints of the Max Planck Institute for Research on Collective Goods Bonn" for the longest part of its existence (some 13,500 pages until 2017/24), and finally to "Discussion Papers of the Max Planck Institute for Research on Collective Goods Bonn" (since 2018/1).

³¹ Preprints 1999/9, 2008/36, 2009/25, 2010/20 and 2012/5.



Graph 1: Annual number of preprints (encircled: ① etc.) and their median length (■) across all three phases of the paper series. The number before “pp.” denotes total page count in each of the phases.

These phases coincided with major turning points in the institute’s evolution: Around the first transition in 2003, it turned from a research group (*Projektgruppe*) into an institute, just as political scientist *Adrienne Héritier*, co-director since April 1998, left the institute in March 2003. The second transition at the start of 2018 came shortly after the accession of co-director *Matthias Sutter* in August 2017.

Consider the development of the number of preprints: Encircled numbers illustrate that peak preprint productivity was in 2008–2010, with another (albeit local) maximum in 2020. The smaller black squares show that median paper length hovered consistently at 30 ± 5 pages,³² but increased by half (to 45 ± 5 pages) at the beginning of the third phase – likely due to the new research group.³³

32 The extremes were rather interesting: Two of the three shortest papers were single-authored by Martin Hellwig (6 pages in 2023/5, 7 pages in 2008/45), both of the longest papers were co-authored by Sebastian Schneider (129 pages in 2023/3, 121 pages in 2023/6). As the document numbers indicate, the shortest and both longest papers were published in 2023 within a sequence of merely 4 papers.

33 Anecdotal evidence: 8 of the 12 longest papers (all 87+ pages) came from authors in the *Sutter* group.

These bibliometric explorations neatly complement a historiography of the institute across its phases of productivity. For instance, its physical relocation marks the difference between preprint 2004/05 (which states the institute's address as Poppelsdorfer Alle 45, 53115 Bonn) and preprint 2004/06 (which mentions Kurt-Schumacher-Str. 10, D-53113 Bonn). Interestingly, even three years after the institute was incorporated (and relocated), preprints still cited its URL as <http://www.mpp-rdg.mpg.de>: Its subsequent (and current) URL <http://www.coll.mpg.de> was not mentioned until preprint 2006/11.

As a final illustration, the preprints contain qualitative data such as the emblem used on their title pages: Starting with the Max Planck Minerva, the series quickly switched to its own oak leaf emblem across four different color schemes, before returning in late 2020 to the Minerva emblem. Here the paper series comes full circle, as tides may have changed in the Max Planck Society's corporate design rules.



Graph 2: All six emblems used for the series' title pages. From left to right, the emblems were last used, respectively, in papers 1998/04, 2005/21, 2014/18, 2019/11, and 2020/24. The right-most emblem is currently still in use.

II. Topics: 588 titles in 2 languages

In order to explore the research topics covered throughout the paper series, I first removed the 13 duplicates identified earlier – seven of which were shorter³⁴ and six longer³⁵ than the successive papers replacing them.

34 By length difference: Papers 2020/26 (45% of 2023/12), 2021/14 (83% of 2024/2), 2023/4 (91% of 2024/8), 2009/25 (94% of 2011/18), 2010/2 (96% of 2015/12), 2010/42 (96% of 2013/23).

35 By length difference: 2007/14 (238% of 2015/4), 2021/6 (230% of 2023/6), 2010/20 (207% of 2011/19), 2017/18 (143% of 2019/9), 2020/17 (135% of 2023/3), 2019/14 (116% of 2020/20), 2005/23 (113% of 2007/2).

I then coded the language of all titles of the remaining 588 papers, using a naïve classification based on nine common stop words from both German³⁶ and English,³⁷ and hand-coding the 68 titles containing none of them. In order to visualize trends in the title vocabulary, I then produced word clouds (where each keyword's font size scales with its frequency) for each of four time periods of six or seven years each.³⁸

The 2×4 roster of word clouds illustrates several shifts in the institute's focus (see next page, graph 3).³⁹

Firstly, the right-hand column exhibits far greater density and font size variation than the left-hand column, resulting from a greater number of English papers in each of the four periods (53, 186, 150, and 108) than German titles in any of them (30, 44, 11, and 7, respectively). Greater sample size means greater variation between frequency extremes, hence the right-hand words vary more in size than the left-hand ones, where most words only appear once.⁴⁰

Secondly, there is an unmistakable evolution in vocabulary, away from applied and policy issues (1998–2003: *produktverantwortung, umweltpolitik, european, governance*; 2004–2010: *wettbewerbspolitik, public, information, taxation*) towards more basic research (2011–2017: *grundlagenforschung, punishment, information, experiment/al*, 2018–2024: *discussion, behavior, preferences, experiment*). Consider the prominence of *experiments*: Barely visible in the first phase, inflected in the second (“experimental”), prominent in both inflections in the third, and clearly dominant in the fourth. Just seven title words appear at least twice in each of the four eras: *public* (32 titles overall), *law* (32), *social* (26), *european* (19), *regulation* (12), *political* (11), and *court* (8). These terms represent the solid core of the institute's topics since its inception.

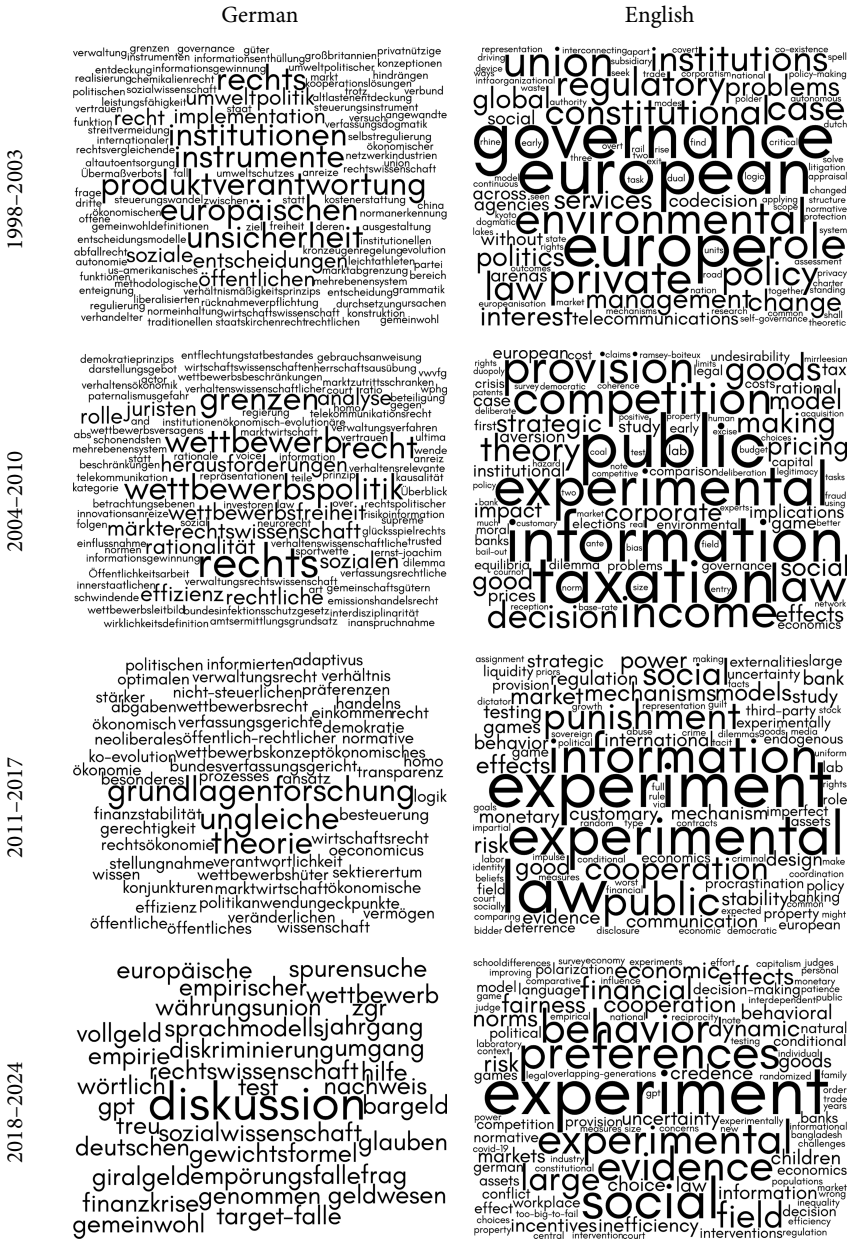
36 Specifically: als, das, der, des, die, eine, und, von, zur.

37 Specifically: a, and, for, not, of, on, the, to, with.

38 I thank my colleague *Rima-Maria Rahal* for this suggestion.

39 Graphs created using *free online wordcloud generator* at www.wordclouds.com with direction “Horizontal”, drawing method “Circles”, min. word size 5, word margin 2, restricted word number.

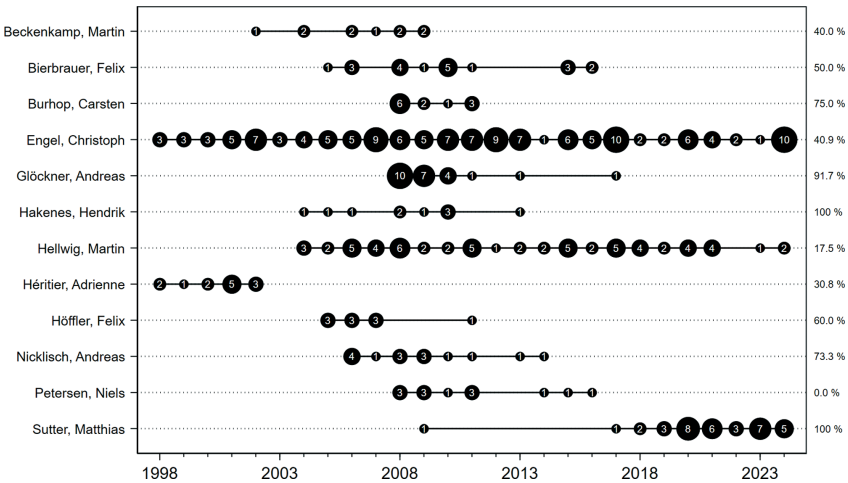
40 In fact, with just 7 and 11 German texts in the latter two periods, they even fell short of 100 words to plot: Their clouds contain 56 title words in the third, and 33 in the last period.



Graph 3: Word clouds of up to 100 most frequent words in paper titles in German (left) or English (right)

III. People: 372 authors in 39 clusters

Lastly, I explored which authors contributed to the series, either alone (in 313 papers, or 53.2 % of the series) or jointly (in 275 papers with 2.7 co-authors on average). Treating all 372 distinct (co-)authors equally, I found that they contributed 2.8 papers on average, and just a dozen authored at least ten papers, shown in the following string-of-beads graph.



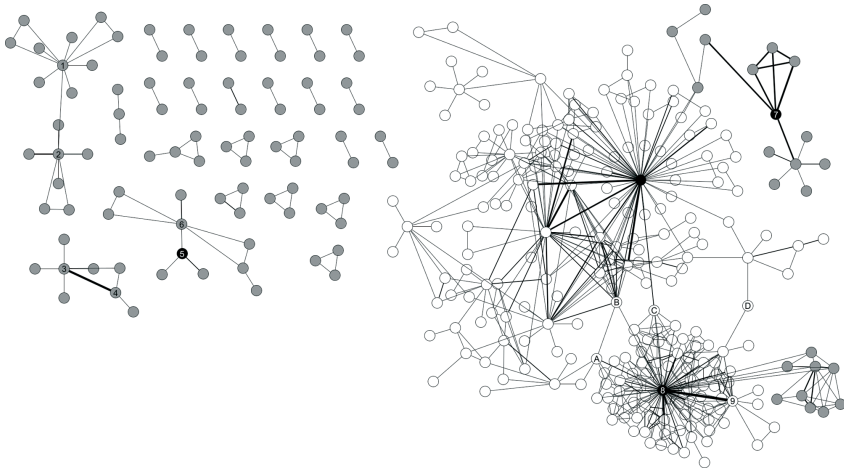
Graph 4: Number of papers per year of the most prolific authors in the series, with their overall co-authorship ratio plotted on the right

The graph reveals that *Christoph Engel* was the series' most prolific author by far, both in absolute terms (137 papers) and relative to the duration of his tenure (5.1 papers p.a.). His three co-directors (*Hellwig*, *Héritier*, and *Sutter*) also contributed papers throughout their tenure, with *Sutter* even once earlier (in 2009), and *Hellwig* still long after he retired (in 2017). Among the directors, co-authorship rates vary starkly: While *Hellwig* co-authored less than one-fifth of his papers, *Engel* co-authored more than two-fifths, and *Sutter* every single paper. Setting directors aside, a junior group leader (*Glöckner*) was the most prolific contributor (24 papers in total), with a record 10 papers in 2008 alone, the year just after *Christoph Engel* had first set this record.

Another author set a remarkable record: All of *Niels Petersen's* 13

papers in the series were single-authored. Including his runner-up, *Carl Christian von Weizsäcker* (6 single-authored), there were just 30 other contributors who authored in the series exclusively by themselves; none of the others contributed more than two papers.

All other authors, who collaborated at least once, can be visualized as nodes in a network where edges represent co-authorships:



Graph 5: Network of co-authors where edge width scales with frequency of co-authorship between adjacent nodes, black nodes mark institute directors, and grey nodes highlight clusters discussed in subsequent text⁴¹

Consider first the right-hand side of the graph: It contains three black nodes, marking directors of the institute. Remarkably, no paper in the series had ever been jointly authored by two directors, which creates easily distinguishable sub-clusters:

At the top right, *Martin Hellwig* (#7) is at the center of a rather stable coauthorship pattern (recognizable as thicker lines) with his coauthors *Gizatulina*, *DeMarzo*, *Pfleiderer*, *Admati*, and *Bierbrauer* – with each of whom he shares three or four papers. Bottom right, we find *Matthias Sutter* (#8) centering a tight sub-cluster of 85 authors – almost a “hairball” in network analysis lingo. It is attached to the remainder of the graph

41 Graph created using yEd 3.23.1 with Layout setting “Organic”.

through only four connectors: *Francesco Feri* (A), *Bernd Irlenbusch* (B), *Bettina Rockenbach* (C), and *Eugenio Verrina* (D). *Sutter* also anchors a smaller cluster of nine authors even further right (in gray) through his coauthors *Walzl*, *Kirchler*, *Stefan*, and *J. Huber*, and he is part of the most stable coauthorship pattern in the entire graph, comprising of six collaborations with *Sebastian Schneider* (#9).

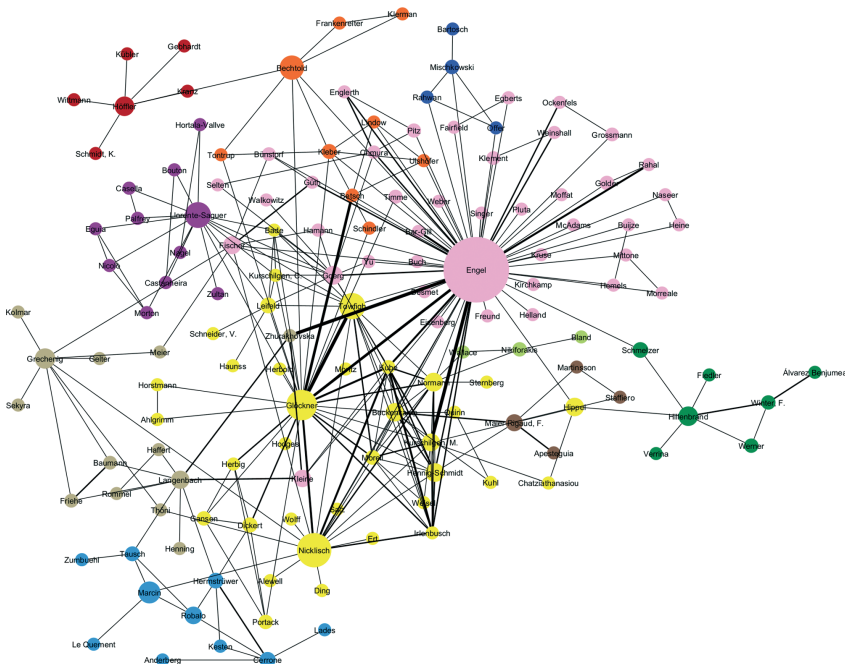
An equally strong collaboration is located at the opposite corner, in the lower left: Another six-fold edge⁴² between *Hendrik Hakenes* (#3) and *Isabel Schnabel* (#4), which forms the backbone of one of 25 self-contained mini clusters in the left half of the graph (in gray). Each of these clusters contains 2–16 authors (or 86 authors in total), who are disconnected from the remainder of the graph. At the top left, the biggest of these clusters centers around *Christian Traxler* (#1) and *Carsten Burhop* (#2), and the third somewhat larger cluster contains a black node, marking director *Adrienne Héritier* (#5), whose co-author *Christoph Knill* (#6) was most central in her cluster.

I next turn to the yet-unexplored main section of the major cluster on the right-hand side, centered around the fourth black node – unsurprisingly, *Christoph Engel*. To put this part of the graph into sharper relief, I removed all grayed clusters and cut the sparse connections at A–D. This left 143 authors, or 38% of the full graph. On these nodes, I ran an edge-betweenness clustering algorithm,⁴³ which resulted in 11 clusters of between 3 and 41 authors. These clusters are color-coded in the following graph and their nodes resized to scale with their betweenness centrality. In other words: The bigger a node, the more shortest paths between all pairs of nodes run through it.

This graph lets us quickly distinguish the ten authors with the greatest centrality within the Engel sub-network. In ascending order: *Felix Höffler*, *Adrian Hillenbrand*, *Kristoffel Grechenig*, *Isabel Marcin*, *Stefan Bechtold*, *Aniol Llorente-Saguer*, *Emanuel Towfigh*, *Andreas Glöckner*, *Andreas Nicklisch*, and – most centrally – *Christoph Engel*.

42 The only other edges of at least five collaborations are between Glöckner/Towfigh, Engel/M. Kurschilgen and Engel/Zhurakhovska.

43 *Girvan/Newman*, Community structure in social and biological networks, PNAS 99 (2002), 7821–7826. Performed using yEd 3.23.1, Auto-Grouping function with Grouping Type “Natural Clusters” and Quality/Time Ratio 1.0.



Graph 6: Network of co-authors connected to *Christoph Engel*, where node size represents betweenness centrality.

Given *Engel's* centrality, we may borrow a famous parameter from the bibliometry of mathematics (the “Erdős number”)⁴⁴ to measure *Engel's* collaborative distance (E) to every other author in the series.⁴⁵ We find that 24% of the entire network (62 authors) directly collaborated with *Engel* at least once ($E = 1$), while 20% of the network (50 authors) and 45% of the network (115 authors) had *Engel numbers* of 2 and 3, respectively. Just five authors in the entire dataset maximised collaborative distance with $E = 5$,⁴⁶ meaning that at least one of their co-authors co-authored

44 See https://en.wikipedia.org/wiki/Erd%C5%91s_number.

45 Except for those who were disconnected from the graph altogether, i.e., 31 exclusive single authors, as well as 86 authors in the self-contained mini-clusters on the left hand side of graph 5.

46 Namely: *Rafael Aigner*, *Pierre C. Boyer*, *Lydia Mechtenberg*, *Nick Netzer*, and *Marco Sahn*.

with someone who co-authored at least once with a co-author of *Christoph Engel*.

D. Discussion/Outlook

Given the limited space allotted to this contribution, I cannot conduct further explorations which might have been interesting. For instance, I did not attempt to manually complete (and utilize) the missing keyword or abstract data (see *supra* II.2), I did not use any of the full text files for which URLs had been collected (see *supra* II.3), I did not match preprints with published research to determine how many preprints resulted in publication, or how much they changed in between. These are just some of the areas that might be explored using the new dataset. Given that retired academics famously have too much time on their hands, I hereby present this dataset as a gift and (hopefully inspiring) pastime to the retiree.⁴⁷ Happy birthday belatedly, *Christoph Engel*!

⁴⁷ Available from Zenodo: Preprints & Discussion Papers of the Max Planck Institute for Research on Collective Goods Bonn (1998–2024) (<https://dx.doi.org/10.5281/zenodo.15479677>).

