Stanford Computational Antitrust

ARTICLE

Competition in the Press: A Computational Analysis of U.S.–German Newspaper Discourse (1870–1945)

Anselm Küsters*

Abstract. What do people mean when they talk about competition? This study maps newspaper discourse on competition in the United States and Germany from 1870 to 1945 using a corpus of approximately 1.1 million digitized articles. Through keyword analysis, dynamic topic modelling, and cross-lingual semantic embeddings, it tracks shifts in themes and tones related to competition throughout industrialization, interwar reforms, and wartime economies. The findings reveal a shift in German competition discourse from civic-commercial vocabulary in the 1870s to organized performance, order, and martial imagery by the 1930s. In contrast, American discourse remained anchored to institutional and economic vocabulary throughout this period. These divergent semantic trajectories are consistent with the different legal paths taken by both countries and suggest that competition narratives are culturally embedded in ways that can hinder transatlantic antitrust coordination and computational antitrust enforcement.

KEYWORDS: Competition; Antitrust; Narrative Economics; Computational Legal History; Cross-Lingual Embeddings; Dynamic Topic Modeling

JEL NOS: K21, L40, N40, D83, C55

* Max Planck Institute for Legal History and Legal Theory, Frankfurt am Main, and Centre for European Policy (cep), Berlin, Germany, kuesters@lhlt.mpg.de. ORCID: 0000-0002-1521-7611.

I. Introduction

What ideas, goals, and mental models lie behind the simple word 'competition'? This question has acquired new urgency. Digitalization has disrupted market structures and geopolitical power on a global scale.[2] Competition law is now tasked with a broader range of objectives beyond allocative efficiency, including calls for labor market protection, reducing inequality, green industrial policy, and democratic resilience.[3] These demands extend from the Neo-Brandeis Movement in the United States[4] to an increasingly multi-purpose approach in Europe.[5] However, the success of these new frameworks will depend on how they are communicated, prioritized,[6] and ultimately enforced – a relationship between regulatory design and social outcome involving far more variables than formal legal analysis typically captures.[7]

One variable that receives almost no attention is the prevailing competition discourse itself, perhaps because it is difficult to quantify. As Ezrachi has argued, competition rules exhibit sponge-like characteristics, absorbing national peculiarities and remaining susceptible to intellectual and regulatory capture.[8] Furthermore, as Shiller's concept of narrative economics suggests, the stories societies tell about markets influence behavior and policy in ways that are

---

[2] JACQUES CRÉMER, YVES-ALEXANDRE DE MONTJOYE & HEIKE SCHWEITZER, COMPETITION POLICY FOR A DIGITAL ERA (2019); ANSELM KÜSTERS, GESTALTUNG DES EU-WETTBEWERBSRECHTS IM DIGITALEN ZEITALTER: EIN QUANTITATIVER UND QUALITATIVER VERGLEICH VON KONSULTATIONSVERFAHREN, EXPERTENBERICHT UND JÜNGSTEN REFORMVORHABEN (1 ed. 2022); ARIEL EZRACHI & MAURICE E. STUCKE, HOW BIG-TECH BARONS SMASH INNOVATION AND HOW TO STRIKE BACK (1 ed. 2022); Patrick Barwise & Leo Watkins, *The Evolution of Digital Dominance: How and Why We Got to GAFA*, *in* DIGITAL DOMINANCE: THE POWER OF GOOGLE, AMAZON, FACEBOOK, AND APPLE 21 (Martin Moore & Damian Tambini eds., 2018).

[3] REBECCA GIBLIN & CORY DOCTOROW, CHOKEPOINT CAPITALISM: HOW BIG TECH AND BIG CONTENT CAPTURED CREATIVE LABOR MARKETS AND HOW WE'LL WIN THEM BACK (1 ed. 2022); Maarten Pieter Schinkel & Leonard Treuren, *Green Antitrust: Friendly Fire in the Fight against Climate Change* in CLIMATE CHANGE & ENVIRONMENTAL SUSTAINABILITY (Simon Holmes, Dirk Middelschute & Martjin Snoep eds., 2021).

[4] Lina Khan, *The New Brandeis Movement: America's Antimonopoly Debate*, 9 J. EUR. COMPET. LAW. PRA. 131, 131 (2018); TIM WU, THE CURSE OF BIGNESS: ANTITRUST IN THE NEW GILDED AGE (1 ed. 2018).

[5] Mariateresa Maggiolino, *The Multi-Purpose Approach and Article 102 TFEU*, *in* FUTURE OF NEO-BRANDEIS MOVEMENT (Thibault Schrepel & Anouk van der Veer eds., 2024).

[6] Or Brook, *Priority Setting as a Double-Edged Sword: How Modernization Strengthened the Role of Public Policy*, 16 J. COMPET. LAW. ECON. 435, 435 (2020) (arguing that prioritization choices by competition authorities function as a policy instrument that shapes the goals of antitrust enforcement).

[7] Since Rafael La Porta, Florencio Lopez-de-Silanes & Andrei Shleifer, *The Economic Consequences of Legal Origins*, 46 J. ECON. LIT. 285, 285-332 (2008) proposed the legal origins theory, scholars have tested various relationships empirically. See, e.g., Sotiris Karkalakos, *The Economic Consequences of Legal Framework*, 45 STATUTE L. REVIEW (2024).

[8] Ariel Ezrachi, *Sponge*, 5 J. ANTITRUST ENFORC. 49, 49-75 (2017).

often overlooked.[9] For instance, historical scholarship has shown that the decision to prosecute international cartels under the Sherman Act, culminating in the 1945 Alcoa ruling, was shaped in part by the perceived connection between cartelized economic power and the rise of European fascism. The discourse surrounding those prosecutions framed market competition explicitly as a pre-condition for democratic governance.[10] This paper takes these observations seriously by following what might be called a narrative antitrust approach.

Understanding the demands and perceptions of competition at different points in time and in different contexts is a worthwhile endeavor.[11] Though it may seem unrelated to antitrust practice at first, examining how historical competition narratives have evolved and diverged across countries can illustrate the variety of demands placed on competition during periods of rapid change, and might even help explain why policymakers, regulators, and observers sometimes talk past each other in international antitrust and trade disputes.[12] This topic is also of interest beyond the field of intellectual history because of the recent trend of using computational methods to enforce the law.[13] However, machine-learning screening tools and large language models (LLMs) trained on U.S. text data may systematically fail to flag or summarize the same conduct when it appears in German or EU corpora framed around different concepts. In short, the semantic tradition embedded in the training data shapes what a model can recognize as harmful. This point is discussed in more depth in Section IV.

More specifically, this paper examines how competition was conceptualized and publicly debated during previous periods of rapid technological advancement in the 19th and 20th centuries, when national competition laws were

---

[9] ROBERT J. SHILLER, NARRATIVE ECONOMICS: HOW STORIES GO VIRAL & DRIVE MAJOR ECONOMIC EVENTS (1 ed. 2019); MICHAEL W.M. ROOS & MATTHIAS RECCIUS, NARRATIVES IN ECONOMICS (1 ed. 2021)

[10] Laura Phillips-Sawyer, *Jurisdiction beyond Our Borders: United States v. Alcoa and the Extraterritorial Reach of American Antitrust, 1909–1945*, in ANTIMONOPOLY AND AMERICAN DEMOCRACY 278, 280–282 (Daniel A. Crane & William J. Novak eds., 2024) (documenting how concerns about European fascism and cartelization drove the Roosevelt administration's decision to extend and revive antitrust prosecution both domestically and internationally).

[11] Just as the characterization of the "consumer" in the competition law domain depends on the context and particular competition law process, the goals and assumptions attached to "beneficial" competition may vary. *See* Albertina Albors-Llorens & Alison Jones, *The Images of the 'Consumer' in EU Competition Law*, in THE IMAGES OF THE CONSUMER IN EU LAW: LEGISLATION, FREE MOVEMENT AND COMPETITION LAW 43 (Dorota Leczykiewicz & Stephen Weatherill eds., 2016). For an economic perspective on this question, see Bruna Ingrao & Claudio Sardoni, *Images of Competition and Their Impact on Modern Macroeconomics*, 27 EUR. J. HIST. ECON. THOU. 500, 500-522 (2020).

[12] Kenneth G. Dau-Schmidt & Carmen L. Brun, *Lost in Translation: The Economic Analysis of Law in the United States and Europe*, 44 COLUM. J. TRANSNAT'L L. 602, 602-621 (2006).

[13] Thibault Schrepel, *Computational Antitrust: An Introduction and Research Agenda*, 1 STAN. COMPUT. ANTITRUST 1, 1-15 (2021).

first proposed with distinctive goals in mind.[14] The Second Industrial Revolution began around 1870 and brought significant changes to production and business organization, including the widespread use of steam power, electricity, and mass production techniques.[15] The United States and Germany, two leading industrial nations at the time, offer a helpful contrast for studying how competition can mean different things to people operating within different institutional and cultural settings. While the U.S. passed a general prohibition on combinations and conspiracies in restraint of trade in 1890 and gradually strengthened it through public enforcement,[16] Germany permitted and eventually mandated cartels, considering organized competition an instrument of industrial policy (for a historical overview, see Section II below).

The period from 1870 to 1945 is relevant because it encompasses the emergence of the first competition-related statutes in both countries and the full arc of German cartel policy, from voluntary association to compulsion. This makes it possible to observe media discourse before and after key legislation was passed. The period ends before American occupation authorities began reexporting antitrust ideas to Germany, thereby altering the very discourse that this paper seeks to study.[17] Two further practical considerations shaped the selection. For both countries, relevant text data exist in a large-scale, digitized form that is accessible for computational research (details in Section II). Furthermore, interpreting the rhetorical register of historical competition discourse requires linguistic and cultural familiarity that automated translation cannot substitute. Thus, the choice also reflects the author's fluency in English and German.

---

[14] For the parallels between the first Industrial Revolution and today's AI age, see ERIK BRYNJOLFSSON & ANDREW MCAFEE, THE SECOND MACHINE AGE: WORK, PROGRESS, AND PROSPERITY IN A TIME OF BRILLIANT TECHNOLOGIES (1st ed. 2014).

[15] The year 1870 is a conventional starting point. It corresponds to the beginning of the Second Industrial Revolution, as outlined in standard economic history literature. It also predates the first competition-related statutes in both countries, enabling an analysis of pre-legislative discourse, DAVID S. LANDES, THE UNBOUND PROMETHEUS: TECHNOLOGICAL CHANGE AND INDUSTRIAL DEVELOPMENT IN WESTERN EUROPE FROM 1750 TO THE PRESENT (1st ed. 1969).

[16] RUDOLPH PERITZ, COMPETITION POLICY IN AMERICA. HISTORY, RHETORIC, LAW (2 ed. 2000); WILLIAM LETWIN, LAW AND ECONOMIC POLICY IN AMERICA: THE EVOLUTION OF THE SHERMAN ANTITRUST ACT (1 ed. 1981).

[17] Notably, the American occupation authorities played a pivotal role in introducing antitrust concepts to Germany and influencing the European Coal and Steel Community (ECSC) Treaty. This framing does not imply that European actors were passive recipients. As Djelic and others have documented, the reception of American antitrust concepts in postwar Germany was selective and contested. It was also filtered through existing ordoliberal traditions. However, the convergence produced by this process makes the prewar divergence studied here all the more significant as a baseline. LISA MURACH-BRAND, ANTITRUST AUF DEUTSCH: DER EINFLUSS DER AMERIKANISCHEN ALLIERTEN AUF DAS GESETZ GEGEN WETTBEWERBSBESCHRÄNKUNGEN (GWB) NACH 1945 (1st ed. 2004); Marie-Laure Djelic, *Does Europe Mean Americanization? The Case of Competition*, 6 COMPET. CHANG. 233, 233-260 (2002).

How can we determine what people in the U.S. and Germany thought about competition during these turbulent times? This question is important for modern enforcement because the assumptions embedded in legal language – what constitutes harmful competition, which forms of coordination are acceptable, what remedies are proportionate, and so on – reflect, at least partly, prior public understandings. These understandings were shaped, in turn, by debates surrounding competition law during their formative decades. As this paper argues at the end, those understandings might be quietly reproduced in algorithmic or LLM-based enforcement tools trained on language-specific corpora that represent one semantic tradition rather than another. To recover dominant discourses about competition during the Second Industrial Revolution, newspaper articles are a key source.

Unlike today's fragmented media landscape, 19th-century newspapers were the primary means of disseminating information and shaping public opinion in both countries.[18] The Industrial Revolution transformed the newspaper industry. Greater profitability drove aesthetic and editorial innovations that made economic and political coverage more prominent.[19] Comparing U.S. and German newspapers from this period thus provides insight into how different legal frameworks, economic conditions, and cultural contexts shaped the meaning of competition[20] during a technological upheaval that disrupted markets and information flows in a way that bears more than a passing resemblance to our own time. Hopefully, scholars with access to other languages and archival sources will expand upon this paper's framework, particularly with French and British sources.

Methodologically, a large-scale discourse analysis of competition is best achieved through computational text analysis.[21] While previous studies have examined the historical context of competition laws[22] and the impact of

[18] RICHARD LEE KAPLAN, POLITICS AND THE AMERICAN PRESS: THE RISE OF OBJECTIVITY, 1865-1920 (1st ed. 2002); MAREK JANSEN, DER NACHRICHTENMARKT IN DER INDUSTRIELLEN REVOLUTION (1st ed. 2022).

[19] Lino Wehrheim, Janos Borst-Graetz, Bernhard Liebl, Manuel Burghardt & Mark Spoerer, *More than a Feeling. Introducing an NLP-Based Media Sentiment Index for the Berlin Stock Exchange, 1872–1930*, 58 HIST. METHODS 139, 139-159 (2025).

[20] For instance, legal scholars have noted that there seem to be strong path dependencies in U.S. and German competition law, for instance regarding how to legally frame the conduct of individual undertakings, Heike Schweitzer, *Parallels and Differences in the Attitudes Towards Single-Firm Conduct: What Are the Reasons? The History, Interpretation and Underlying Principles of Sec. 2 Sherman Act and Art. 82 EC*, EUI LAW (2007), https://hdl.handle.net/1814/7626.

[21] FRANCO MORETTI, DISTANT READING (1st ed. 2013); Michael A. Livermore & Daniel N. Rockmore, *Distant Reading the Law*, in LAW AS DATA 3, 3-19 (Michael A. Livermore & Daniel N. Rockmore eds., 2019).

[22] David J. Gerber, *The Origins of European Competition Law in Fin-de-Siecle Austria*, 36 AM. J. LEGAL HIST. 405, 405-440 (1992); DAVID J. GERBER, LAW AND COMPETITION IN TWENTIETH-CENTURY EUROPE: PROTECTING PROMETHEUS (1st ed. 1998); Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J.L.& ECON. 7, 7-48 (1966).

technological change on industrial development,[23] they have not, to the best of the author's knowledge, used natural language processing (NLP) to analyze large bodies of historical texts.[24] This paper employs three NLP methods in a deliberate sequence. Each method addresses a limitation of the one before it. Keyword frequencies establish the basic vocabulary of competition discourse, but they cannot distinguish thematic context. For example, the word competition appears in sporting fixtures and cartel proceedings alike. Dynamic topic modeling resolves this issue by grouping words into co-occurrence clusters and tracking how those clusters shift across decades. It does, however, compress meaning into a small number of discrete themes. The third method, cross-lingual semantic embeddings, addresses that limitation. An embedding model represents each word or document as a dense, low-dimensional vector in a continuous mathematical space, where geometric proximity corresponds to semantic similarity: terms that consistently appear in similar contexts are placed close together. This differs fundamentally from a traditional bag-of-words (BoW) model, which represents texts as high-dimensional sparse vectors of raw word counts and therefore treats any two words as equally distinct regardless of how they are actually used. This matters especially for normative vocabulary. For instance, a BoW model treats fair competition and unfair competition as nearly identical, because they share a root token. An embedding model places them in distinct regions of the semantic space, because the two phrases consistently co-occur with different arguments and legal outcomes: one with protective and coordinative language, the other with prosecutorial and prohibitive language. This paper's methodology is designed to exploit precisely this capacity to recover normative distinctions at scale, across two languages, and over seven decades.

The paper is organized into three sections. Section II anchors the analysis in the historical context and describes the data foundation. Section III is the main part of the paper, presenting a computational analysis of competition discourse in the U.S. and Germany over time. This section uses increasingly sophisticated methods: first, keywords, then, estimation of a dynamic topic model and finally, cross-lingual embeddings. Section IV offers tentative interpretations based on close readings of examples and arguments from specialized sports literature in German historiography, elaborating on the German peculiarities identified in the empirical results. It also connects semantic results to legal debates and draws lessons for international antitrust coordination and the computational enforcement of the law. Section V concludes.

---

[23] LANDES, *supra* note 15.

[24] For introductions, see MATTHEW LEE JOCKERS, MACROANALYSIS: DIGITAL METHODS AND LITERARY HISTORY (1st ed. 2013); TED UNDERWOOD, DISTANT HORIZONS: DIGITAL EVIDENCE AND LITERARY CHANGE (1st ed. 2019); JUSTIN GRIMMER, MARGARET E. ROBERTS & BRANDON M. STEWART, TEXT AS DATA: A NEW FRAMEWORK FOR MACHINE LEARNING AND THE SOCIAL SCIENCES (1st ed. 2022).

II. Historical Background and Data

Before turning to the data sources and computational results, it is useful to provide a brief historical contextualization of the antitrust discourse in the United States and Germany. At the beginning of the 1870s, neither country had a modern competition statute. Markets in both countries were shaped by contract doctrine, guild traditions, and emerging industrialist associations.

From the outset, the first legislative responses to mounting public concern were philosophically distinct. The Sherman Act of 1890,[25] the Clayton Act of 1914,[26] and the Federal Trade Commission Act of 1914[27] established competition as a matter of public law and subject to permanent federal oversight in the U.S., reflecting widespread alarm at the market power of the great trusts.[28]

Germany's trajectory was different.[29] The 1896 Unfair Competition Act regulated commercial behavior rather than market structure. Germany's decisive shift toward managed competition occurred during the hyperinflation of 1923 and the subsequent Cartel Ordinance. This legislation recognized cartels as legitimate economic organizations subject to public oversight rather than prohibition.[30] Further pro-cartel measures followed under the Nazi regime.[31]

These legislative milestones are the historical and interpretive anchors for the subsequent computational patterns revealed by the data. Figure 1 shows that, based on the corpus of this paper (for the underlying data sources, see below), media attention to competition consistently spiked at these moments in both countries. In the U.S., coverage increased after the Sherman Act of 1890, with a notable uptick following the Clayton Act of 1914. In Germany, media attention rose after the Unfair Competition Act of 1909 and again following the Ordinance against the Misuse of Economic Power of 1923. A pronounced surge appeared in the early 1930s, coinciding with the Nazi regime's interventions enabling compulsory cartels

[25] Sherman Antitrust Act of 1890, ch. 647, 26 Stat. 209 (codified as amended at 15 U.S.C. §§ 1–7).

[26] Clayton Antitrust Act of 1914, ch. 323, 38 Stat. 730 (codified as amended at 15 U.S.C. §§ 12–27).

[27] Federal Trade Commission Act of 1914, ch. 311, 38 Stat. 717 (codified as amended at 15 U.S.C. §§ 41–58).

[28] RUDOLPH PERITZ, COMPETITION POLICY IN AMERICA. HISTORY, RHETORIC, LAW (2 ed. 2000); LETWIN, *supra* note 16.

[29] Gerber, *Origins*, *supra* note 22; GERBER, LAW AND COMPETITION, *supra* note 22.

[30] Thomas Jovović, *Deutschland und die Kartelle – eine unendliche Geschichte*, 53 JAHRBUCH FÜR WIRTSCHAFTSGESCHICHTE / ECONOMIC HISTORY YEARBOOK 237, 237-273 (2012).

[31] Sigrid Quack & Marie-Laure Djelic, *Adaptation, Recombination and Reinforcement: The Story of Antitrust and Competition Law in Germany and Europe*, *in* BEYOND CONTINUITY: INSTITUTIONAL CHANGE IN ADVANCED POLITICAL ECONOMIES 255 (Wolfgang Streeck & Kathleen Ann Thelen eds., 2005).

and tighter state oversight.[32] Keeping this chronology in mind is essential for interpreting the qualitative and quantitative results that follow.

The corpus underlying this study and Figure 1 were compiled to capture the public discourse surrounding the development of domestic competition laws. For the U.S., the articles were drawn from the *American Stories* dataset, which was compiled by Dell et al.[33] This dataset comprises nearly 20 million scans of articles from the Library of Congress's *Chronicling America* collection, which is in the public domain.[34] The dataset was filtered to retain articles published between 1870 and 1945, i.e., the temporal scope set by the theoretical motivation of this paper, rather than by the boundaries of the dataset itself, which extend beyond these endpoints. Retrieval was performed using a single keyword search for the term *competition* with case-insensitive matching and word-boundary constraints. The keyword *antitrust* was not used as a supplementary retrieval term, despite its increasing currency from approximately 1903 onward. This choice was deliberate on two grounds. First, the paper's core interest is in competition discourse broadly construed, not in antitrust debate as a discrete political genre; using antitrust would have introduced a post-1890 selection bias that skews the pre-legislative sample. Second, and more fundamentally, no functional German equivalent of antitrust developed during the study period, so including it for the U.S. sub-corpus would have compromised cross-national comparability in precisely the years, i.e. the early 1900s through the 1910s, when divergence is most analytically significant. This means that articles that frame competitive conduct exclusively in the language of trusts, monopoly, or Sherman Act, without using the word competition may be underrepresented in the post-1903 U.S. sub-corpus. To assess the scale of this gap, a supplementary retrieval using the broader keyword set {*competition*, *antitrust*, *monopoly*, *cartel*} was performed. It yielded 358,276 U.S. articles, compared with 257,625 from the competition-only search. The topic model and embedding results reported in Section III are therefore best understood as capturing competition discourse in its widest semantic register rather than as a complete record of antitrust coverage.

For the German portion of the corpus, newspaper pages were obtained from the *Deutsche Digitale Bibliothek* (DDB) newspaper portal, which covers approximately 4.5 million pages from 1671 to 1950. The keyword *Wettbewerb* (competition) was applied to the same date range and duplicates were removed,

---

[32] Jovović, *supra* note 30.

[33] Melissa Dell, Jacob Carlson, Tom Bryan, Emily Silcock, Abhishek Arora, Zejiang Shen, Luca D'Amico-Wong, Quan Le, Pablo Querubin & Leander Heldring, *American Stories: A Large-Scale Structured Text Dataset of Historical U.S. Newspapers*, ARXIV (2023), https://doi.org/10.48550/arXiv.2308.12477.

[34] Earlier versions of this collection had unreliable article segmentation and poor optical character recognition (OCR) quality. Dell et al. solved these issues with a deep learning pipeline that uses a transformer-based layout detection model to identify article boundaries on scanned pages and a custom OCR engine fine-tuned on historical newspaper typefaces. This improved the text quality significantly compared to the legacy output of the Library of Congress.

yielding 834,731 pages. Because the DDB Application Programming Interface (API) delivers full pages rather than individual articles, these records are not directly comparable to the 257,625 individual articles in the U.S. sub-corpus. After applying the sentence-window extraction procedure described below, 834,731 page-level records were reduced to competition-proximate text passages; the resulting unit count of extracted passages (one per page-record) is used in all subsequent analyses. For transparency, all longitudinal frequency analyses are normalized by the total record count per country-year, so that volume differences between the two archives do not drive cross-national comparisons.

To adjust for the structural difference between the two archives, each German page was reduced to a passage related to competition by keeping every sentence containing the keyword *Wettbewerb* and the ten sentences before and after it. Any overlapping windows were merged, and one text was kept per record. This brought the average German text length to 196 words, close to the average length of a 361-word U.S. article. Further text cleaning to reduce noise from Optical Character Recognition was applied to both sub-corpora following established practice.[35]

To what extent does press discourse mentioning competition concern what is now called antitrust rather than everyday rivalry in sports, civic life, or commerce? To answer this question, two country-specific dictionaries of explicit antitrust terminology were compiled and matched against the corpus.[36] The results were normalized to reflect newspaper growth. The measure is necessarily approximate, but the proportions it yields are informative.[37] Still, it should be noted that neither sub-corpus is a balanced sample of the full press spectrum. The *Chronicling America* collection leans toward regional and small-town papers and is not weighted by circulation. Similarly, the DDB API reflects German archival digitization priorities. Drawing on research conducted for comparable newspaper digitization projects, one might expect these biases to underrepresent working-class and labor-movement framings of competition. Labor and socialist papers, which in both countries tended to frame market rivalry as exploitation rather than as a

[35] Following best practices in NLP, GRIMMER, ROBERTS & STEWART, *supra* note 24, chapters 3 and 4.

[36] The German dictionary comprises 20 terms, including historical orthographic variants, and covers the following concepts: *Konkurrenz* (and its historical variant *Concurrenz*), *Kartell* and compounds (*Kartellgesetz*, *Kartellamt*, *Kartellbildung*, *Kartellrecht*, *Kartellverfahren*, *Kartellabkommen*, *Kartellaufsicht*, *Preiskartell*), *Monopol* and compounds (*Monopole*, *Monopolbildung*, *Monopolstellung*), *Missbrauch*, *Preisabsprachen*, *Preiskontrolle*, and *Konkurrenzschutz*. The U.S. dictionary comprises 17 pattern entries covering 13 distinct concepts: antitrust, Sherman Act, Clayton Act, FTC (and FTC Act), cartel(s), monopoly/monopolies/monopolization, price fixing (and pricefixing), restraint of trade, competition law, antimonopoly, trust-busting, and trust(s). Both dictionaries use case-insensitive whole-word matching with word-boundary anchors; multi-word expressions use whitespace-tolerant regex patterns.

[37] For example, including "trust(s)" improves U.S. coverage of pre-1914 rhetoric, but it risks capturing commercial "trust" usage.

productive social mechanism, have often been disproportionately excluded from early digitization programs, which prioritized commercially successful metropolitan titles with intact archival holdings.[38] As a result, framings of competition as wage suppression, labor degradation, or class conflict are, therefore, likely understated in both sub-corpora. This limitation should be borne in mind when interpreting the normative valence of the results.

On average, antitrust-proximate articles account for 9.35 percent of U.S. articles and 9.23 percent of German articles across the full period; a surprisingly similar baseline, considering the countries' divergent legal frameworks. U.S. coverage spikes sharply during the so-called Great Merger Movement and its legal denouement, reaching 21.9 percent in 1899 and 22.9 percent in 1911. This is consistent with the consolidation wave documented by Lamoreaux, as well as the *Standard Oil* and *American Tobacco* decisions.[39] By contrast, Germany maintains a broad plateau of 9-11 percent from the 1890s through the early 1910s without comparable peaks.[40]

Both series converge during the First and Second World Wars, reaching their lowest point as discourse about economic competition gives way to vocabulary of coordination and mobilization. This is a pattern worth noting. The wartime suppression of antitrust rhetoric might be a recurring dynamic in which a national emergency reframes market rivalry as a collective effort to be managed rather than a process to be protected. This dynamic has an obvious contemporary resonance, as geopolitical competition and AI-driven industrial policy are generating pressure once again to subordinate antitrust to strategic priorities.

The series diverge sharply again in the early New Deal and Nazi years. U.S. antitrust attention recovers modestly while German antitrust discourse falls below 5 percent as economic categories are displaced by the vocabulary of national mobilization and racial struggle.

---

[38] Kaspar Beelen, Jon Lawrence, Daniel C.S. Wilson & David Beavan, *Bias and Representativeness in Digitized Newspaper Collections: Introducing the Environmental Scan*, 38 DIGIT. SCHOLARSH. HUM. 1, 5-12 (2023) (distinguishing between representation bias, i.e. the divergence between the digitized sample and the known newspaper landscape, and historical bias, i.e. the footprint of social and historical inequalities embedded in the source material itself; the paper demonstrates, using the JISC corpus of Victorian British provincial newspapers, that digitized collections can systematically over-represent politically aligned and more expensive titles while under-representing neutral and cheaper papers).

[39] The Great Merger Movement (approximately 1895–1904) was a wave of horizontal consolidations in U.S. manufacturing that concentrated substantial market power in sectors including steel, oil refining, meatpacking, and tobacco. Its "legal denouement" refers to the Supreme Court's landmark decisions in *United States v. Standard Oil Co.*, 221 U.S. 1 (1911), and *United States v. American Tobacco Co.*, 221 U.S. 106 (1911), both of which applied the rule of reason to break up consolidated industrial combinations formed during the merger wave, NAOMI R. LAMOREAUX, THE GREAT MERGER MOVEMENT IN AMERICAN BUSINESS, 1895-1904 (1 ed. 1985); LETWIN, *supra* note 16.

[40] This plateau is punctuated by an outlier in 1881 (26.2 percent on a very small denominator) and a gentle post-1914 slide that deepens after 1933.

III. Computational Analysis of Competition Discourse

This section shifts from counting to close reading. First, it provides a keyword-based reading of the national competition discourses captured in the newspaper corpus (Sub-section A). Building on that, it estimates a dynamic topic model that allows the thematic composition of competition discourse to change over time (Sub-section B). Finally, cross-lingual embeddings are used to compare meanings across both sub-corpora (Sub-section C).

Throughout the analysis, it is worth bearing in mind that historical newspapers are generally noisy, multi-genre texts.[41] To clarify the text, the first two sections of the computational analysis thus operate on a part-of-speech (POS)-filtered corpus that singles out nouns and adjectives – the carriers of topics in news prose – while preserving the orthographic variants typical of the period.[42] In line with the motivation set out in the Introduction, the goal of this computational analysis is not to determine legal correctness retroactively, but rather to recover the language available at the time through which readers encountered competition.

III.1 Keyword Approach

Which words appeared most often alongside competition in the national press? Frequency lists by country contrast ubiquitous vocabulary with characteristic rhetorical markers. The patterns in Tables 1 and 2 are computed from the full competition-keyword corpus, i.e. all articles retrieved on the basis of the keyword competition (U.S.) or Wettbewerb (Germany), rather than from the narrower antitrust-proximate sub-corpus used to track legal coverage in Section II. This choice reflects the section's goal of recovering the broadest semantic environment of competition, including sporting, civic, and commercial registers, not only explicitly antitrust content. In each table, the left pair of columns ranks the most common words by raw frequency, or how often they appear across all articles. The right pair of columns ranks words by their mean TF-IDF (term frequency–inverse document frequency) score. TF-IDF identifies terms that appear more frequently in competition articles than in the broader newspaper corpus as a whole, i.e. words that are characteristic of competition coverage specifically, not merely common in the

[41] Lino Wehrheim, Bernhard Liebl & Manuel Burghardt, *Extracting Textual Data from Historical Newspaper Scans and Its Challenges for "Guerilla-Projects"*, REGENSBURG ECONOMIC AND SOCIAL HISTORY DISCUSSION PAPER SERIES (December 2, 2022), http://doi.org/10.5283/epub.53259 (describing practical challenges of OCR pipeline construction for historical newspaper corpora, including segmentation errors and typography-driven noise); Brian Beach & W. Walker Hanlon, *Historical Newspaper Data: A Researcher's Guide and Toolkit*, 90 EXPLOR. ECON. HIST. 101541, 101553-101557 (2022) (providing a systematic overview of quality assessment procedures for digitized newspaper archives).

[42] In practice, I create country-specific stop lists, normalize text, and retain only nouns, adjectives, and proper nouns based on the corresponding spaCy models. Documents with fewer than five meaningful tokens are discarded.

press at large. High TF-IDF scores function as a more informative signal of what competition meant in each national context than raw frequency alone, which is dominated by general-purpose vocabulary.

Rank	Most frequent words (raw count)	Count (absolute)	Most distinctive words (TF-IDF)	Count (mean)
1	competition	305,094	man	0.0207
2	man	192,117	price	0.0181
3	state	146,670	company	0.0177
4	company	131,513	state	0.0172
5	business	121,691	team	0.0170
6	price	119,697	business	0.0162
7	american	119,222	american	0.0161
8	city	116,624	club	0.0156
9	country	114,151	city	0.0153
10	ing	105,107	country	0.0142
11	high	104,993	high	0.0138
12	team	93,410	game	0.0132
13	club	91,645	united	0.0123
14	united	90,822	ing	0.0122
15	large	90,524	school	0.0121
16	work	84,002	line	0.0117
17	ton	83,684	trade	0.0117
18	second	80,179	large	0.0116
19	people	80,121	week	0.0115
20	line	80,074	work	0.0115

Table 1. Word frequency and TF-IDF rankings, U.S. competition corpus, 1870-1945. Left columns: twenty most frequent nouns and adjectives by raw count. Right columns: twenty most distinctive terms by mean TF-IDF score, weighted against the full U.S. newspaper corpus.

Unsurprisingly, the U.S. frequency table (Table 1) begins with the word competition (305,094 counts), followed by general public affairs nouns such as man, state, company, business, price, American, and city. The TF-IDF ranking emphasizes a more distinctive word list. Alongside price, company, state, and business, we see team, club, game, and week.[43]

[43] The "ing" token is a statistical artifact and is likely a byproduct of OCR and hyphenation.

This is the first piece of evidence to suggest that athletics constitute a significant part of the discourse on competition. The presence of terms such as trade and state in the TF-IDF list is consistent with the classical preoccupation of U.S. antitrust with restraints on trade.[44] Overall, the picture is of a press where business vocabulary co-exists with the language of leagues and fixtures.

Rank	Most frequent words (raw count)	Count (absolute)	Most distinctive words (TF-IDF)	Count (mean)
1	deutsch [German]	1,012,835	deutsch [German]	0.0356
2	wettbewerb [competition]	975,761	preis [price]	0.0206
3	preis [price]	392,432	meter [meter]	0.0180
4	deutschland [Germany]	335,739	berlin [Berlin]	0.0166
5	berlin [Berlin]	307,204	deutschland [Germany]	0.0165
6	letzter [last]	291,559	letzter [last]	0.0137
7	land [country]	280,050	land [country]	0.0137
8	meter [meter]	272,671	mark [mark]	0.0114
9	mann [man]	189,306	mann [man]	0.0110
10	frage [ask]	181,344	verein [association]	0.0108
11	hoch [high]	179,630	sonntag [Sunday]	0.0107
12	stadt [city]	169,460	stadt [city]	0.0106
13	sonntag [Sunday]	168,241	iſt [Iſt]	0.0104
14	mark [mark]	165,862	punkt [point]	0.0104
15	arbeit [work]	162,747	frage [ask]	0.0101
16	verein [association]	161,737	hoch [high]	0.0100
17	punkt [point]	159,113	arbeit [work]	0.0098
18	staat [state]	154,618	industrie [industry]	0.0094
19	reich [rich]	153,483	staat [state]	0.0092
20	kampf [battle]	147,434	england [England]	0.0091

Table 2. Word frequency and TF-IDF rankings, German competition corpus, 1870-1945. Left columns: twenty most frequent nouns and adjectives by raw count. Right columns: twenty most distinctive terms by mean TF-IDF score, weighted against the full German newspaper corpus. English translations in brackets.

[44] HERBERT HOVENKAMP, ENTERPRISE AND AMERICAN LAW, 1836-1937 (1 ed. 1991) (examining how nineteenth-century American courts and legislators constructed the legal framework for competitive enterprise, including the statutory origins of the restraint-of-trade doctrine); LETWIN, *supra* note 16.

The German keyword lists (Table 2) show a different mix. In terms of raw frequency, *Deutsch* (German) and *Wettbewerb* (competition) are dominant, followed by *Preis* (price), *Deutschland* (Germany), *Berlin*, and civic or sporting terms such as *Meter* (meter), *Sonntag* (Sunday), *Verein* (club), *Punkt* (point), and *Kampf* (fight). In the TF-IDF table, economic and institutional nouns – such as *Preis* (price), *Mark* (German mark, i.e. the currency at the time), *Industrie* (industry), and *Staat* (state) – remain prominent. In particular, the role of *Preis* and *Mark* hints at a discourse where competition is tied to price debates/regulation and currency stability, i.e., themes that loomed large in interwar German debates about cartels and administered markets.

From 1870 until the end of World War II, German courts generally treated cartel contracts as enforceable agreements. At the same time, *Verein*, *Meter*, and *Sonntag* anchor a substantial sporting strand tied to Germany's dense association system (*Vereinswesen*) and weekend athletics, a point further explored below. National markers, i.e. *deutsch*, *Deutschland*, and *England*, are more prevalent than in the U.S. lists, suggesting that contests (industrial or athletic) were often framed in terms of national competition. This nationalistic tone aligns with contemporary reporting styles and foreshadows later semantic trends during the Nazi period.

Taken together, the keyword approach, complemented with close reading of illustrative example snippets, reveals two facts that shape the subsequent analysis. First, the legal and economic registers of the two countries diverge at the lexical level. The U.S. keyword profile emphasizes trade, business, restraint, and price; a cluster consistent with the statutory vocabulary of the Sherman Antitrust Act of 1890, which prohibited combinations and conspiracies in restraint of trade. Indeed, Progressive Era reformers and prosecutors built their cases around entry barriers, market structure, and the institutional conditions for price competition. The rhetorical target was concentrated on private power expressed through contract and combination, not disorderly or unfair commercial conduct as such. According to the frequency counts, this structural vocabulary proved durable. It survived the transition from trust-busting to the rule-of-reason era, through New Deal price regulation, and into the postwar period. The keyword data captures that stability at the lexical level. In contrast, German keywords reflect a tradition in which competition requires organization and oversight rather than unimpeded market forces. Second, athletic vocabulary systematically appears in both corpora, but its relationship to economic vocabulary differs. In the U.S., sports terms appear alongside market terms within a shared discourse of institutional competition. In contrast, in the German list, sporting terms are denser and increasingly separated from regulatory language.

As the dynamic topic model will demonstrate in the next section, this separation widens dramatically from the 1920s onward. During this period, economic competition discourse in Germany gradually contracts, while the language of performance and order expands to fill its place.

III.2 Dynamic Topic Modeling

A topic model is a statistical method that automatically discovers recurring thematic clusters across a large body of texts.[45] For example, given tens of thousands of articles, it identifies groups of words that tend to appear together and treats each group as a latent topic. Words such as price, cartel, and trade tend to co-occur in one set of articles, while team, club, and score cluster in another. The model identifies these groupings without being told what they are in advance. In standard latent Dirichlet allocation (LDA), the topics are static; the model estimates a single set of themes assumed to remain consistent throughout the analyzed period.

Historical newspapers violate this assumption because the vocabulary and concerns associated with a given theme evolve as institutions, laws, and political realities change. Therefore, this section uses a dynamic topic model (DTM) instead, which extends LDA by allowing each topic's vocabulary to gradually change over time.[46] One can intuitively think of this setup as a topic model with institutional memory, where themes in the 1890s organically grow out of those in the 1880s rather than being estimated in isolation. After setting the parameters and allowing the model to converge,[47] each topic can be summarized by its decade-specific top words.

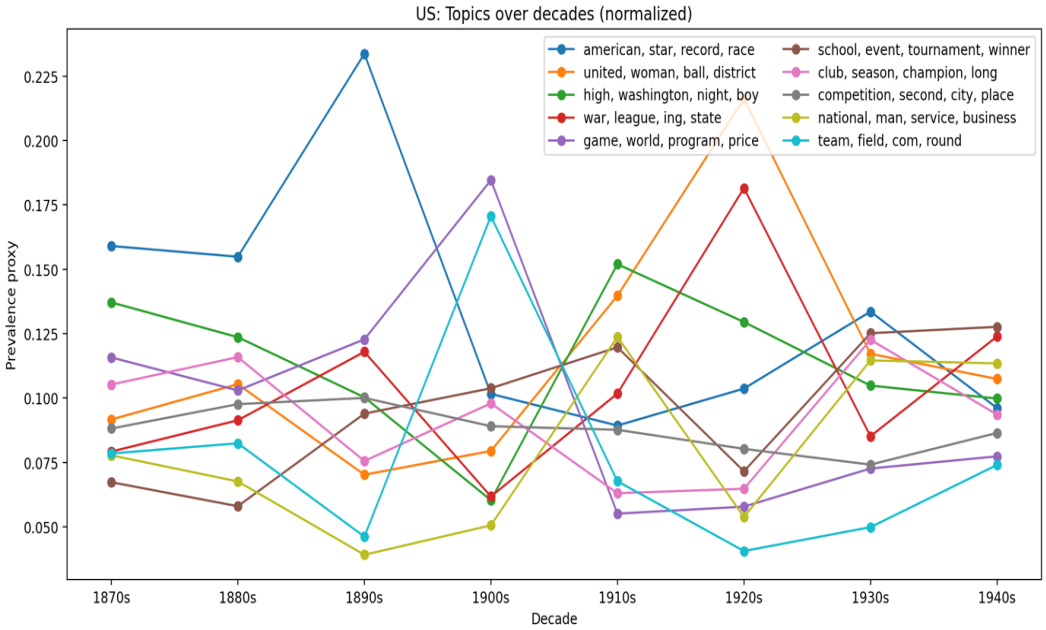
---

[45] David M. Blei, Andrew Y. Ng & Michael I. Jordan, *Latent Dirichlet Allocation*, 3 J. MACH. LEARN. RES. 993, 993-1022 (2003) (introducing the generative probabilistic model for topic discovery and deriving the variational EM inference procedure); Thomas L. Griffiths & Mark Steyvers, *Finding Scientific Topics*, 101 PROC. NATL. ACAD. SCI. U.S.A. 5228, 5228-5235 (2004) (applying the topic model approach and explaining how topics can be interpreted in practice).

[46] David M. Blei & John D. Lafferty, *Dynamic Topic Models*, *in* ICML '06: PROCEEDINGS OF THE 23RD INTERNATIONAL CONFERENCE ON MACHINE LEARNING 113, 114 (2006); David Mimno, *Computational Historiography: Data Mining in a Century of Classics Journals*, 5 J. COMPUT. CULT. HERIT. 1, 13 (2012).

[47] Specifically, all articles are divided into ten-year periods (1870s to 1940s) and separate DTMs are applied to the U.S. and German sub-corpora using the *Tomotopy* Python package with 10 topics per country, (again) PoS-filtered tokens, and 150 training iterations. The topic counts are deliberately small (k = 10) to favor broad themes over specific subgenres. Sensitivity checks with a higher k value further split sports but leave economic clusters and their trajectories intact. The decade bins smooth year-to-year volatility and avoid overfitting to small denominators in the early decades. However, this comes at the cost of blurring short-lived spikes.

Figures 2 and 3 plot the relative share of the competition corpus occupied by each thematic cluster for each decade. A rising topic's characteristic vocabulary



appears more frequently alongside the keyword competition, while a declining topic's bundle of words recedes from the discourse. The Appendix provides the underlying word lists for readers who wish to examine the model output in greater detail.

Figure 1. Relative share of each estimated topic in the competition corpus, by decade (1870s-1940s). Each line represents one of the ten estimated topics, labeled by its most characteristic words from the 1940s.

The German DTM reveals three identifiable transitions, each of which tracks a major institutional shift. In the early decades, the dominant register linked rivalry to civic administration and skilled craft. Terms such as trade, law, association, and traffic frame competition as a regulated social activity rather than an anonymous market process. This vocabulary is consistent with guild traditions and the absence of any general prohibition on cartels. A second phase, spanning from the late 1880s to the 1910s, is shaped by the emergence of unfair competition laws. Price, legal, and organizational terms cluster together, reflecting the consolidation of the Law Against Unfair Competition[48] and the cartel coordination debates that followed shortly after. The third and most striking transition begins in the 1920s and accelerates sharply after 1933. Athletic and associational vocabulary, such as Sunday, meter, club, and champion, gains weight across several topics. Meanwhile, explicitly economic competition terms recede and are ultimately crowded out by terms with unambiguous National Socialist resonance. Rivalry is still present in the data but has been re-metaphorized as organized performance, collective rank, and martial struggle. Overall, the German DTM traces an arc from

[48] Until 1894, German courts resisted controls on "unfair competition" because they were considered incompatible with the freedom to trade. The first statutory foothold came in 1894, followed by the 1896 Law to Combat Unfair Competition, and culminating in the comprehensive 1909 Law Against Unfair Competition mentioned in the Introduction.

market/price/law to performance/title/order, a semantic shift that the subsequent embedding analysis quantifies more precisely.

The U.S. DTM tells a different structural story: one of semantic stability rather than displacement. From the 1870s through the 1910s, American newspaper articles consistently cluster around the vocabulary of transportation infrastructure, state regulation, and market access, such as railroads, commissions, tariffs, and rights-of-way. By the Progressive Era, administrative markers had become firmly embedded alongside competition, reflecting the institutionalization of trust-busting as a recognizable policy genre with its own established vocabulary. Sports vocabulary has grown throughout this period, but the model clearly separates it into distinct categories. It never absorbs the policy register the way the German data show. Throughout the New Deal era, competition remained close to government, law, and federal commissions. Therefore, the U.S. DTM yields the mirror image of the German finding. Rather than dissolving into performance and order, competition in the American press became progressively institutionalized, acquiring a regulatory vocabulary that proved resilient across political cycles once established in the 1890s.

DTM outputs identify the dominant thematic clusters in each decade. Close reading of the highest-scoring articles in each topic-decade cell verifies that the clusters capture meaningful discourse rather than statistical artifacts and recovers the specific rhetorical registers that vocabulary lists can only approximate. This two-step procedure, which involves first estimating a generative model and then reading the documents that the model ranks highest within each topic category,[49] is standard practice in Digital Humanities research.[50] Computational metrics surface candidates that would be impossible to identify by hand at this scale, and close reading then confirms that the model assignments are semantically coherent and historically interpretable. The examples below should therefore be seen as textual evidence that grounds each computational pattern described above in public argument and concrete sources.

[49] The procedure works as follows. For each decade-topic cell, a vocabulary was constructed from the forty most frequent words in the model's topic-decade distribution. Each article in the corpus was then scored by computing the count of token-level matches between its POS-filtered tokens and that decade-topic vocabulary, i.e. how many of the article's retained nouns and adjectives appear in the topic's characteristic word list for that decade. This match count is the "overlap" score. The 5 articles with the highest overlap scores in each topic-decade cell were read in full. This ensures that the close-reading examples are selected on the basis of the model's output rather than researcher judgment.

[50] Jordan Boyd-Graber, David Mimno & David Newman, *Care and Feeding of Topic Models: Problems, Diagnostics, and Improvements*, *in* HANDBOOK OF MIXED MEMBERSHIP MODELS AND THEIR APPLICATIONS 259 (Edoardo M. Airoldi et al. eds., 2014); Jordan Boyd-Graber, Yuening Hu & David Mimno, *Applications of Topic Models*, 11 FOUNDATIONS AND TRENDS IN INFORMATION RETRIEVAL 143, 170 (2017).

III.2.A Germany: From Civic Order to Organized Performance

The first glimpse into German discussions about competition opens in the slowly industrializing world of the 1870s. During this decade, association life was repeatedly presented as an appropriate way to discipline markets through craft and character. For instance, a Württemberg arts-and-crafts association publicly urged industrial education to focus on the future representatives of German industrial diligence (*künftigen Vertreter des deutschen Gewerbefleißes*) and called for statutes that cultivated competition by raising standards and disseminating techniques, revealing contemporary skepticism toward laissez-faire competition, or Manchesterism, as it was known then. A further text from this decade that is highly-ranked by the model is a report from April 1879 on a lecture by Ernst Carl Ludwig Bratuscheck (1837–1883), a German classical philologist and philosopher. The lecture was held at the *Volksbildungsverein* in Gießen. Contrasting the household economy, which is held together by duty and love, with the national economy, which is degraded by egoism, Bratuscheck warned that unrestricted competition unleashes a struggle for existence that only the economically stronger can win. His remedy was not price control, but civic education, so economic freedom could advance alongside culture. This text exemplifies the anti-Manchester register that the DTM attributes to the decade: moralized, associational, and skeptical of self-correcting markets.

This stance (competition yes, unfairness no) became the rhetorical scaffolding for the 1909 Law Against Unfair Competition through a process of discursive consolidation. Consumer associations and trade journals documented deceptive labeling and adulteration, and mounting public pressure forced politicians, and later courts, to treat these practices as legally cognizable wrongs rather than merely poor commercial judgment.[51] The newspaper corpus captures this accumulation across three decades. What begins as a moral complaint in the 1870s becomes organized associational lobbying in the 1890s, as exemplified by the articles on the *Verein gegen Unwesen im Handel und Gewerbe*. These articles called for a statute modeled on Code Civil §1382, which holds perpetrators of unfair competition liable for compensating victims. Significantly, the official drafting materials invoked the same *Kampf ums Dasein* rhetoric found in the newspaper corpus, stating that competition "can at times become a fight with unequal weapons, in which honest trade must come off worse".[52] Previously, the courts had resisted such claims; the *Apollinaris* decision of 1880 rejected the application of §1382 of the Code Civil against unfair competition.[53] However, organized civic pressure, as documented in the Reichstag commission reports, ultimately forced a statutory

[51] HENNING STECHOW, DAS GESETZ ZUR BEKÄMPFUNG DES UNLAUTEREN WETTBEWERBS VOM 27. MAI 1896: ENTSTEHUNGSGESCHICHTE UND WIRKUNG 85, 89, 126, 141, 146–65 (1st ed. 2019).

[52] Own translation based on AXEL BEATER, NACHAHMEN IM WETTBEWERB: EINE RECHTSVERGLEICHENDE UNTERSUCHUNG ZU [PARAGRAPHEN] 1 UWG 43 (1 ed. 1995).

[53] *Id.* at 8–13, 59.

solution. Media discourse reflected and amplified the shift in public expectations that eventually produced the statute.

In late 19th century Germany, fair competition thus referred to a civic-linguistic order based on claims in public markets, rather than the static efficiency of the later Chicago School. It functioned as an order of publicity that stabilized expectations.[54] As a growing number of U.S. students trained at German universities absorbed ideas from the Historical School – influenced by Friedrich List, who advocated protectionism – and returned home, Progressive Era writers also began to reframe rivalry as fair.[55] Similar arguments appear in the other top-ranked articles from this period. Liberal economists acknowledged that unrestricted competition in foodstuffs resulted in systematic adulteration. Tariff commentators reframed import competition as asymmetric dependence requiring state intervention. They treated *Einfuhrzoll* (import duty) not as a restraint on competition, but as an instrument that would restore fair competition between domestic and foreign producers. Industrial policy reports legitimized state subsidies for shipping lines, including the Reich's subsidized East Asia steamship line, as a standard instrument of national commercial strategy. This increasingly blurred the line between competition policy and colonial transport statecraft.

The anchor text for the decades of pre-World War I is a 1910 analysis of the coal syndicate and the steelworks association. It is the highest-ranking article in its decade-topic cell. The article predicted that Prussia would "do everything possible" to ensure the syndicate's survival, explaining that large, integrated firms depended on self-consumption privileges for coal and semi-finished steel. Thus, organization at one level propped up organization at another. This corresponds precisely to empirical findings from economic history: cartel prices in the German steel industry were set below profit-maximizing levels, yet well above competitive benchmarks.[56] Users accepted higher input prices in exchange for preferential terms and reliable supply.[57] Overall, German newspapers normalized competition structured by associations, privileges, and state instruments. In that language, disorderly competition is both inefficient and uncivil.

---

[54] David J. Gerber, *Constitutionalizing the Economy: German Neo-Liberalism, Competition Law and the "New" Europe*, 42 AM. J. COMP. L. 24, 42 (1994); GERBER, *supra* note 22.

[55] Daniel T. Rodgers, ATLANTIC CROSSINGS: SOCIAL POLITICS IN A PROGRESSIVE AGE 76-77, 82-84 and 110 (1 ed. 2010).

[56] Margaret C. Levenstein & Valerie Y. Suslow, *What Determines Cartel Success?*, 44 J. ECON. LITERATURE 43, 81 (2006).

[57] Daniel Barbezat, *Structural Rigidity and the Severity of the German Depression: The AVI and the German Steel Cartels 1925-1932*, 31 EXPLORATIONS ECON. HIST. 479, 483 (1994).

III.2.B United States: Institutional Stability and Structural Vocabulary

The U.S. DTM reveals a different discourse. Throughout, competition remains coupled with state and business topics, whether under trust-busting, the FTC's unfair methods rubric, or wartime price control debates. The model captures a public conversation that repeatedly returns to entry barriers and institutional means to check private power. The dominant register is infrastructural and quasi-regulatory rather than moral.

For instance, from the 1870s to the 1890s, urban transportation coverage framed competition as a matter of franchises and rights-of-way, i.e. competition for the market rather than within it, which explains the transportation vocabulary in Table 1. The strongest example is a New York proceeding regarding the Second Avenue Railway. In this proceeding, commissioners set a rental price for street access and insisted that the company compensate the city for occupying it in a public way. This is an early articulation of scarcity rents and congestion externalities in utility competition, anticipating today's debates about platforms and network access. Additional top-ranked articles from this period document how railroads responded to water transport competition by controlling access to docks and wharves. A 1909 speech by Joseph H. Call of the Municipal Waterways Association condemning the railroads' systematic effort to "acquire the waterfront and secure ownership and control of the docks and wharves" reads like a primer on vertical foreclosure and essential facilities, long before these terms were formalized in doctrine. Other high-ranking articles from the U.S. sub-corpus cover carload rate disputes, municipal coal yards established to undercut local cartels, and interstation distance tariffs. These articles also present competition as a structural outcome to be engineered through franchise allocation and rate regulation rather than a moral condition to be judged.

The anchor text for the post-Sherman Act era, as surfaced by the DTM, is reporting on the Union Pacific–Southern Pacific case from 1909. In this case, the government sought to prove that there had been active competition between Western lines prior to the merger. The government framed stock control as a direct violation of the Sherman Act, arguing that the merger eliminated all competition. By 1921, the FTC complaint against U.S. Steel's Pittsburgh pricing system extended this logic to pricing practices. During the National Recovery Administration period, the government's steel price coordinator condemned collusive base prices as "non-competitive prices, lacking the safeguards for consumers that competition provides", explicitly invoking the procompetitive benchmark. Throughout, the U.S. DTM rhetoric is strikingly non-moralized: it counts lines, stock blocks, and market actors rather than invoking fairness idioms. It portrays market structure, not commercial character, as the primary explanation for the presence or absence of competition.

Taken together, the DTM patterns and close reading examples reveal distinct national grammar of competition. In the U.S., competition is about prices,

entry, terminal access, and the institutions that constrain discrimination or foreclosure. When competition fails, the solution is legal and structural. In Germany, the late 19th-century vocabulary of competition progressively intertwined with association, morality, civic education, and state coordination. By the 1920s and 1930s, it was also associated with sports and organized performance. This shifts the center of gravity from decentralized market rivalry to rank, order, and state management. References for price and industry remain frequent, but the rhetorical frame shifts. Competition becomes an instrument for ordering economic life rather than a spontaneous process requiring protection from private power.

III.3 Embedding Approach

Words that appear in similar contexts tend to have similar meanings. Embedding models operationalize this idea by assigning a numerical address to each word or text in a large coordinate space. Items that share contextual associations are positioned close together, while items that do not share these associations are positioned far apart. The result is a kind of semantic map, where proximity corresponds to semantic similarity. This differs from simply counting how often a word appears, an approach that treats each word in isolation with no consideration for its context, and from topic modeling, which reduces vocabulary to a few broad themes and loses the distinctions that might matter for legal and cultural analysis. Early work in this tradition focused on individual words. The landmark word2vec and GloVe models demonstrated that models trained on large text collections could learn, without human supervision, that the words king and queen are related in the same way as man and woman.[58] Later approaches extended this concept to sentences and documents, enabling comparisons of larger linguistic units that more accurately reflect context. For instance, these representations have been used to model bodies of legal cases.[59] Likewise, economic research uses embeddings to quantify shifts in tone and content across central bank speeches, linking language to market reactions.[60] Historical studies have used diachronic embeddings to trace the evolution of stereotypes and social meanings in public discourse.[61]

---

[58] Tomas Mikolov, Kai Chen, Greg Corrado & Jeffrey Dean, *Efficient Estimation of Word Representations in Vector Space*, 1ST INT. CONF. ON LEARNING REPRESENTATIONS, WORKSHOP TRACK PROC. 1, 2 (2013).

[59] Elliott Ash & Daniel L. Chen, *Case Vectors: Spatial Representations of the Law Using Document Embeddings*, *in* LAW AS DATA 313, 317 (Michael A. Livermore & Daniel N. Rockmore eds., 2019).

[60] Stephen Hansen & Michael McMahon, *Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication*, 99 J. INT'L ECON. S114, S114-S115 (2016); Martin Baumgärtner & Johannes Zahner, *Whatever It Takes to Understand a Central Banker - Embedding Their Words Using Neural Networks*, 157 J. INT'L ECON. 104101, 104101-104102 (2025).

[61] Nikhil Garg, Londa Schiebinger, Dan Jurfafsky & James Zou, *Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes*, 115 PROC. NATL. ACAD. SCI. U.S.A. E3635, E3635, E3639-40 (2018).

For the question of what the term competition meant in newspapers in the U.S. and Germany from 1870 to 1945, embeddings provide a common metric for comparing meaning across languages and time. In particular, this approach enables us to examine whether references to competition are more closely aligned with economic-administrative discourse or a combative, warfare-like register that overlaps with military and athletic rhetoric.

Operationally, this section encodes newspaper texts from the joint corpus using a compact cross-lingual transformer model known as *Paraphrase-Multilingual-MiniLM-L12-v2*.[62] For this model, cross-lingual training ensures that differences in proximity reflect substantive semantic alignment, rather than translation *per se*.[63] This allows for a direct comparison of German and English articles.

Again, the core idea behind the embedding approach is to assign a numerical address to every newspaper article in a high-dimensional space. Articles that discuss similar topics end up close together, while articles that discuss different topics end up far apart. Each address, called a vector, is computed by passing the full text of the article through the cross-lingual model and averaging the numerical representations it produces for each word, a standard technique known as mean pooling. Importantly, the original article text is used instead of the grammatically filtered versions used in the keyword and topic modeling sections because the full-sentence context gives individual words their meaning. To keep the analysis computationally manageable, a balanced sample (319,770 articles) was created, approximately 160,000 from each country, to ensure proportional representation of each ten-year period from 1870 to 1945.[64]

The analysis then asks the following question. In articles that mention competition, does the surrounding vocabulary lean more toward economic and administrative language or war and contest language? To answer this question, two

[62] The *Paraphrase-Multilingual-MiniLM-L12-v2* model was chosen because it translates both English and German into a shared vector space. This enables direct semantic comparisons without the need for intermediate translations. Furthermore, with 384 dimensions and 12 transformer layers, the model strikes a balance between representational richness and computational tractability. Importantly, the model is trained on paraphrase pairs rather than translation pairs. This means it encodes semantic proximity rather than translational equivalence. This property is exactly what is needed to determine whether German and English articles about competition inhabit the same conceptual neighborhood. Nils Reimers & Iryna Gurevych, *Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks*, *in* PROCEEDINGS OF THE 2019 CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING AND THE 9TH INTERNATIONAL JOINT CONFERENCE ON NATURAL LANGUAGE PROCESSING 3982, 3982-85 (Sebastian Padó and Ruihong Hang eds., 2019).

[63] Mikel Artetxe & Holger Schwenk, *Massively Multilingual Sentence Embeddings for Zero-Shot Cross-Lingual Transfer and Beyond*, 7 TRANSACTIONS OF THE ASS'N FOR COMPUTATIONAL LINGUISTICS 597, 597-598 (2019).

[64] Specifically, I construct a balanced sample with decade-level stratification, capping it at 200,000 items per country, and then equalizing it on the smaller side.

semantic reference points, called bundle vectors, are constructed for each country.[65] Each bundle is simply the average of the numerical representations of a short list of semantically related words. One list covers economic and administrative vocabulary, the other list covers war and contest vocabulary.[66] This technique of constructing reference points by averaging word vectors is established in the literature.[67]

For each country-decade combination, the average position of all articles in that cell, known as the centroid, is calculated. Intuitively, the centroid represents the typical semantic neighborhood in which competition appears in a given country during a given decade. Then, the distance between the centroid and each reference bundle is measured using cosine similarity, a standard metric that assesses the angle between two vectors in high-dimensional space.[68] A small angle means the two points are semantically close, a right angle means they are not very similar, and a wide angle means they are semantically opposed.

Finally, a single summary measure, the semantic tilt, is calculated for each country-decade as the cosine similarity with the economic bundle minus the cosine similarity with the war/contest bundle. A positive tilt indicates that competition discourse during that period is more economically oriented, while a negative tilt suggests a lean toward combative or martial vocabulary. Figure 2 plots the cosine series and resulting tilt for each country across the full period, revealing the diverging national trajectories.

---

[65] This practice is similar to constructing dimension vectors in cultural semantics. Austin C. Kozlowski, Matt Taddy & James A. Evans, *The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings*, 84 AM. SOCIO. REV. 905, 911-19 (2019).

[66] Economic-administrative bundle for Germany: *Verlust, Handel, Kapital, Industrie, Verkauf, Produktion, Markt, Wirtschaft, Unternehmen, Preis, Firma, Absatz, Fabrik, Gewinn, Geschäft*. Economic-administrative bundle for U.S.: *loss, trade, capital, industry, sale, production, market, economy, company, price, firm, sales, commerce, enterprise, factory, profit*. War/contest bundle for Germany: *Krieg, Kampf, Feind, Angriff, Verteidigung, Schlacht, Eroberung, Soldaten, Armee, Militär, Waffen, Strategie, Taktik, Sieg, Niederlage, Macht, Herrschaft, Widerstand, Offensive, Defensive, Front, Kämpfer, Bedrohung, Zerstörung*. War/contest bundle for U.S.: *war, battle, enemy, attack, defense, victory, defeat, soldiers, army, military, weapons, strategy, tactics, combat, fight, power, domination, resistance, offensive, defensive, front, fighter, threat, destruction*. Note: These lists are fixed lexicons used only to generate the corresponding mean-embedding "bundle" vectors, *i.e.* they are not dictionaries of all possible synonyms. In Python, the reference bundles are all lowercased.

[67] Sanjeev Arora, Yingyu Liang & Tengyu Ma, *A Simple but Tough-to-Beat Baseline for Sentence Embeddings*, 5 INT'L CONF ON LEARNING REPRESENTATIONS 171, 171-186 (2017).

[68] Again, this approach of taking the difference in cosines is related to existing literature. Aylin Caliskan, Joanna J. Bryson & Arvind Narayanan, *Semantics Derived Automatically from Language Corpora Contain Human-like Biases*, 356 SCIENCE 183, 183-186 (2017); Kozlowski, Taddy & Evans, *supra* note 65.

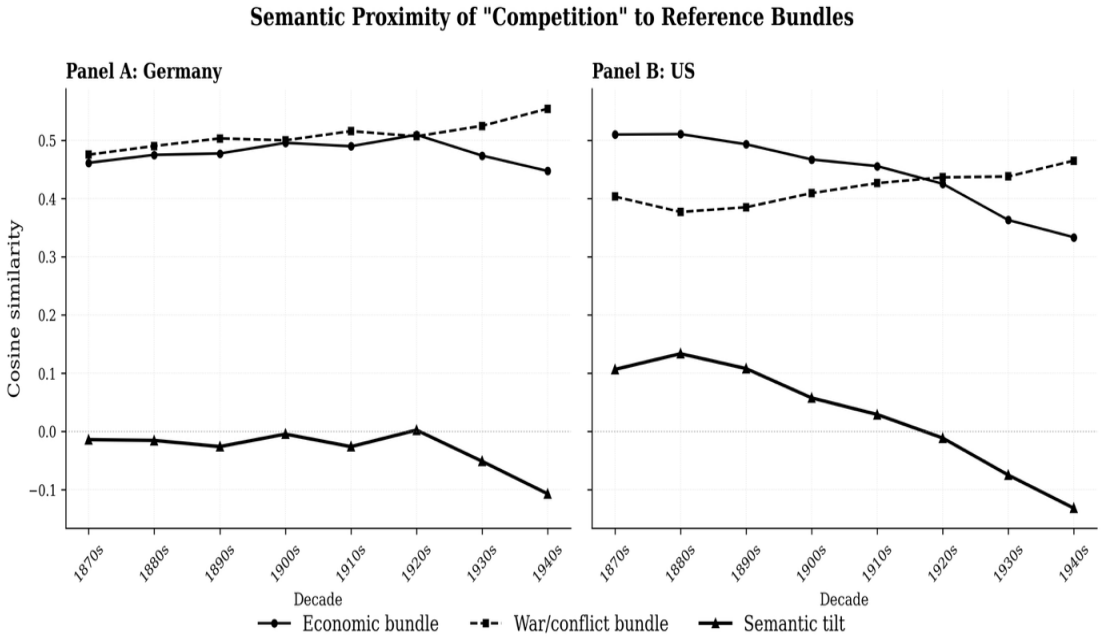


Figure 2. Semantic positioning of competition-related newspaper articles via cross-lingual embeddings. Panel A: German discourse. Panel B: U.S. discourse. Each panel plots three time series: (1) cosine similarity between competition passages and the economic-administrative bundle, (2) cosine similarity to the war and conflict bundle, and (3) the semantic tilt score (economic similarity minus war similarity, triangles). Positive tilt values indicate stronger association with economic vocabulary; negative values indicate stronger association with war or conflict language. Source: approximately 160,000 articles per country, Paraphrase-Multilingual-MiniLM-L12-v2 transformer embeddings.

The computational results confirm diverging semantic trajectories between the two countries (see Figure 2). On the German side, competition discourse exhibits a consistently war-related orientation throughout most of the period. Tilt scores remain negative throughout this period (-0.014 in the 1870s, -0.015 in the 1880s and -0.026 in the 1890s), reflecting the enduring influence of metaphors of struggle from military and athletic contexts in commercial settings. While the gradual decline across these three decades can only be interpreted causally with further research, it is consistent with two concurrent developments in the corpus. First, the Franco-Prussian War of 1870-71 and the subsequent unification of the German Empire likely injected martial vocabulary into public discourse at the precise moment the corpus opens. Newspapers that covered economic subjects in this period did so against a backdrop of military mobilization. Second, the late 1880s and 1890s saw the accelerating growth of German export industry, which German newspapers routinely framed as national commercial struggle against British and American competitors, i.e. the *Konkurrenzkampf* vocabulary documented in the illustrative examples below. Both factors probably pushed competition discourse away from the economic-administrative pole before any domestic competition statute existed to anchor it there. The brief recovery in the 1900s (-0.003), approaching zero, coincides with the consolidation of the Unfair Competition Act of 1909. The 1920s is the only decade in the German series where economic usage outweighs war-related usage, coinciding with the Weimar

stabilization period and the enactment of the first domestic competition statute, the 1923 Cartel Ordinance.

By contrast, the 1910s show no such moderation: the tilt falls back to -0.026, which is roughly the same as in the 1890s. This is consistent with the mobilization vocabulary of the First World War entering the corpus. After 1923, however, the trajectory deteriorates rapidly, with the tilt dropping to -0.051 in the 1930s and -0.107 in the 1940s. This is because National Socialist discourse progressively dissolved market categories into existential-struggle vocabulary, subordinating economic competition to national mobilization.

The U.S. trajectory appears similar but is structurally different. American competition discourse opens with strongly positive tilt scores (+0.107 in the 1870s, +0.134 in the 1880s and +0.108 in the 1890s), indicating that economic and administrative vocabulary dominated the semantic neighborhood of competition from the outset. The contrast with Germany, with its negative tilt at this same moment, is stark. The gap widens further through the 1890s, where the U.S. figure significantly diverges from the German figure. Crucially, this divergence predates the Sherman Antitrust Act of 1890 and the Progressive Era trust-busting campaigns. The implication is that the semantic separation between U.S. and German competition discourse was not primarily a product of different legal regimes. Rather, the legal regimes appear to have been built on pre-existing semantic foundations. American newspaper readers encountered competition in a discursive environment already dominated by market and institutional vocabulary. German newspaper readers encountered the same word in a discursive environment already tilted toward contest and national mobilization. The initial U.S. orientation is sustained during the Progressive Era. The tilt remains clearly positive in the 1900s (+0.058) and 1910s (+0.029), consistent with the amplification of trust-busting rhetoric and the clustering of competition with regulatory and market-structure vocabulary. The first crossing of the zero line occurs in the 1920s (-0.011), reflecting the tentative introduction of cooperative and quasi-military language into public economic debate during the transition to the New Deal. The decline then steepens sharply, reaching -0.075 in the 1930s and -0.132 in the 1940s, driven by full wartime mobilization.

While the convergence in tilt values by the 1940s is superficially striking, it conceals a persistent divergence in what those numbers represent. American wartime competition discourse retains the vocabulary of federal oversight, price coordination boards and institutional regulation; the war frame is applied to an essentially unchanged market-structural conception of competition. By this point, German discourse had substantially replaced the market-structural frame with one of collective struggle and territorial defense, meaning that competition itself, and not merely its regulation, was narrated in martial terms.

To verify which specific words drive this divergence, each term in the two reference bundles was temporarily removed and the similarity score recalculated

without it. A word that causes a large drop when removed is one that was doing substantial work in pulling the typical competition narrative toward the economic or the martial side. The persistent intermingling of sports and military vocabulary in German competition discourse, evident in phrases such as competitive battle (*Konkurrenzkampf*) or defensive struggle (*Abwehrkampf*), reveals a semantic foundation that equates market rivalry with a physical contest requiring coordination and discipline. On the U.S. side, the words doing the heaviest work in pulling American competition discourse away from the economic pole are not combat terms like fight or defeat, which contribute modestly and positively, but only a few words from the organized-sports vocabulary of institutions and outcomes. The potential legal consequences of these narrative traditions will be discussed in the next section.

We can use some concrete examples from the German newspaper dataset to further illustrate and ensure the validity of the embedding findings. As early as 1877, Leipzig's bid for the *Reichsgericht*'s seat was framed as a contest with a trophy ("adorned with the prize of victory [...] from the competition with Berlin", *geschmückt mit dem Siegespreis [...] aus dem Wettbewerb mit Berlin*). An 1879 programmatic piece explicitly equated competition with emulation (*Wetteifer*, literally translatable as "rivalry in striving for excellence") and urged merchants to surpass rivals in quality rather than price, reframing exchange as a noble contest of excellence. In the late 19th and early 20th century, cross-channel trade tensions were repeatedly depicted as a fight for commercial supremacy, a bout-like struggle (*Konkurrenzkampf*, meaning "competitive battle" or "struggle for competitive advantage"). By 1905, Germany's export push was depicted as a wrestling match, a breathless struggle in increasingly difficult competition with foreign markets ("a breathless struggle in an increasingly difficult competition with foreign markets", *atemloses Ringen [im] immer schwerer werdenden Wettbewerb mit dem Ausland*). During the 1909 Reichstag debates, quoted in newspapers, small and mid-sized firms were said to face large firms "as if in a fight" (*wie in einem Kampf*). By 1925, British agitation against German exporters was described as a "a battue, a driven hunt, against German competition" (*Kesseltreiben gegen den deutschen Wettbewerb*), a hunting-sport metaphor that portrayed industry as prey in a coordinated pursuit.

The articles in the German sub-corpus demonstrate that, over the decades, the German press has repeatedly narrated non-sport competition acts using an athletic and military lexicon that transforms economic and institutional contests into races with lanes, matches with rules, and fights that award victory. Race metaphors ("overtaken/outpaced", *überholt*; "into ordered channels/lanes", *in geordnete Bahnen*) position international or sectoral competition as a regulated event that can be managed and judged, thereby implicitly legitimizing policy interventions as rule-setting rather than market-distorting. Prize-and-victory tropes ("palm of victory", *Siegespalme*; "prize of victory", *Siegespreis*) elevate economic outcomes into honorific contests, drawing on the meritocratic imagery familiar to contemporary newspaper readers from sports. Fair-play formulas ("within the

bounds of what is permitted", *im Rahmen des Erlaubten*; "friendly-neighbourly competition"; *freundnachbarlicher Wettbewerb*) domesticate legal disputes by foregrounding permissible bounds. This foreshadows the later "rules of the game" (*Spielregeln des Wettbewerbs*) rhetoric in German and EU competition law.[69] At the same time, contest language shades into combative registers (e.g., "defensive struggle" or "fight of resistance", *Abwehrkampf*). The sports frame overlaps with struggle narratives that justify defensive coordination against rivals perceived overreach. Together, these tropes render competition as a culturally legible, rule-bound game that rewards initiative but depends on lane-keeping, refereeing, and shared norms of fair play. To this day, this framing underwrites the legitimacy of *Ordnungspolitik* ("the politics of order", which is the German tradition of using regulatory frameworks to structure markets rather than prohibit private coordination) in German discourse.[70]

IV. From Newspapers to Doctrine: Semantic Divergence, Legal Tradition, and Computational Enforcement

The embedding results in Section III.3 point to a pattern that neither keyword frequencies nor topic modeling alone could have established. German competition narratives were oriented toward martial and athletic registers from the earliest decades of the study period, before any domestic competition statute existed, and that orientation deepened rather than corrected as legal institutions developed. Sports vocabulary does not simply co-exist alongside economic vocabulary in the German sub-corpus. From the 1920s onward, it progressively displaces it. This section asks what explains that displacement and what follows from it for legal doctrine and computational enforcement. The sociogenesis of this rhetorical trajectory could be explained by how football and soldierly discipline became intertwined in German public life after 1870. A body of specialized literature, mostly authored by sports historians, documents this process. Broader cultural-historical accounts of performance (*Leistung*) as a governing concept remain scarcer.[71]

What follows summarizes the key findings from that literature, sketches how they connect to the legal domain from a narrative antitrust perspective and

[69] ANSELM KÜSTERS, THE MAKING AND UNMAKING OF ORDOLIBERAL LANGUAGE. A DIGITAL CONCEPTUAL HISTORY OF EUROPEAN COMPETITION LAW 76, 359, 423-432 (1st ed. 2023).

[70] On the different dimensions of this discourse, see Karen Horn, *Ordoliberalism: Neither Exclusively German nor an Oddity. A Review Essay of Malte Dold's and Tim Krieger's Ordoliberalism and European Economic Policy: Between Realpolitik and Economic Utopia*, 35 REV. AUSTRIAN ECON. 547, 547-560 (2021).

[71] According to Wolfram Pyta, *Vom Segen zum Fluch? Der Beitrag von Leistung und Wettbewerb zur Karriere des Sports in Deutschland*, *in* WIRTSCHAFTLICHER UND SPORTLICHER WETTBEWERB: FESTSCHRIFT FÜR RAINER GOMMEL ZUM 65. GEBURTSTAG 239, 239-256 (Rainer Gömmel, Markus A. Denzel, & Margarete Wagner-Braun eds., 2009). But see also NINA VERHEYEN, DIE ERFINDUNG DER LEISTUNG (1st ed. 2018).

draws consequences for international enforcement coordination and AI-assisted competition law.

Eggers documents the gradual institutionalization of football in Germany after 1900, the debates about rules, and the incorporation of football into military training before 1914. This created a pipeline that normalized rule-bound trials of strength as a form of civic education.[72] Wartime plays and postwar re-mobilization then popularized this connection. Barracks-derived rituals, martial terminology, and the ideal of tough, disciplined comradeship etched a soldierly template into club culture.[73] This relationship survived long after 1918. By the early Weimar years, football was on its way to becoming a *Volkssport*, with standardized competition formats that newspapers could serialize week after week.[74] Eisenberg's emphasis on the urban or even bourgeoise origins of German sports is certainly valid. However, the diffusion into smaller towns in the 1920s widened the audience for this repertoire and multiplied the local club news that are partly captured in the corpus used in this study.[75] The more readers consumed league tables (i.e. the ranked standings that newspapers serialized week after week), points totals, championship titles, and referee-ordered contests, the more plausible it became to re-define competition as a *Leistungskampf*, i.e. a contest where merit is measured, ranked, and publicly certified.[76]

The village level confirms the same moral economy of achievement. Wittmann's analysis of a representative, small-town football club in the Weimar Republic, whose anniversary booklet documented the club's history and communal standing, revolves around a narrative in which repeated good performances generate sporting success and, consequently, broader social recognition.[77] The chronicle assumes the key logic, whereby performance yields legitimacy, legitimacy yields members, and members yield further performance. This is not an isolated case study. As we saw in the German topic list estimated by the DTM, *Leistung*, *Meisterschaft*, and *Verein* frequently appeared together on the weekend pages. Meanwhile, weekday economic pages continued to debate prices. Thus, the newspapers were essentially broadcasting a pedagogy of rank, discipline, and improvement that could easily travel from athletics to firms and public offices. In that semantic environment, the idea of competition as a fair trial took root.

---

[72] Erik Eggers, *Die Anfänge des Fußballsports in Deutschland: Zur Genese eines Massenphänomens*, in FUßBALL ALS KULTURPHÄNOMEN: KUNST - KULT - KOMMERZ 67, 73–79 (Markwart Herzog & Ulrich von Berg eds., 2002).

[73] Markwart Herzog, *Footballers as Soldiers. Rituals of Masculinity in Twentieth-Century Germany: Physical, Pedagogical, Political, Ethical and Social Aspects*, 43 REV. STADION 250, 250-269 (2019).

[74] Eggers, *supra* note 72, at 79-90.

[75] Christine Eisenberg, *Fußball in Deutschland 1890-1914*, 20 GESCH. GES. 181, 181 (1994).

[76] Pyta, *supra* note 71.

[77] Florian Wittmann, *Eine neue Quelle zur dörflichen Fußballgeschichte in der Weimarer Republik*, 19 SPORTZEITEN 43, 43 (2019).

It is precisely here that the legal doctrine of *Leistungswettbewerb*, typically translated as competition on the merits,[78] comes in as a formalization of the press's routine framing. As the German unfair competition tradition evolved from the 1870s onwards, its general clauses provided an opportunity that jurists in the 1930s, especially Hans Carl Nipperdey and Franz Böhm, could expand into a principled distinction between *Leistungswettbewerb* and *Behinderungswettbewerb*.[79] As the author has reconstructed elsewhere, the *Benrather Tankstellen* case from 1931 solidified this vocabulary. Subsequently, courts condemned "economic rule of force" (*wirtschaftliches Faustrecht*) and valued acts that win customers through better performance rather than by disabling rivals.[80] Böhm's famous definition, that competition is "not a fight man against man, but a race, i.e. the performance of the participants is not taking place as a clash, coming from opposing directions as in a duel, a wrestling match, a preliminary fight, or a war, but in a parallel direction",[81] translated the contemporary sports page into doctrine. The works of Walter Eucken, one of the architects behind Germany's social market economy, later popularized the metaphor: "performance-based competition resembles a race [...] and at the finishing line the consumers should decide".[82] In this framework, the state must guard the rules to ensure that the trial of merit remains legible and that coercion is excluded.

The German embeddings computed above, which pull competition toward performance and sports in the earlier decades, align with this legal rhetoric. These semantic associations of the term matter because the doctrine of *Leistungswettbewerb* identifies an ethic without a definitive test for difficult situations, which leaves discretion to the interpreter.[83]

Compared to the German approach, the U.S. discourse appears more administrative and legal than moral and ethical. This difference is reflected in the different keywords, the DTM's topic structure, and the distinct pattern of the U.S. embeddings. In the newspaper articles surveyed here, the U.S. competition discourse remains tied to commissions, rates, branding, and procurement. When it reaches

[78] The phrase competition on the merits is now firmly established in the case law of Article 102 TFEU and closely corresponds to *Leistungswettbewerb* as it has been developed within the German ordoliberal tradition. However, as I found in my doctoral thesis, it is not always translated in a standardised way in existing EU case law. See KÜSTERS, *supra* note 69.

[79] BEATER, *supra* note 52, at 71-86.

[80] KÜSTERS, *supra* note 69, at 74-78. See also, Franz Hederer & Kim Christian Priemel, *In der Schwebe. Markt, Staat und Wettbewerb in Deutschland zwischen 1918 und 1948*, 313 HIST. Z. 89, 105 (2021).

[81] FRANZ BÖHM, DIE ORDNUNG DER WIRTSCHAFT ALS GESCHICHTLICHE AUFGABE UND RECHTSSCHÖPFERISCHE LEISTUNG 124 (1 ed. 1937).

[82] WALTER EUCKEN, GRUNDSÄTZE DER WIRTSCHAFTSPOLITIK 42 (7th ed. 2004).

[83] BEATER, *supra* note 52, at 71-86.

for fairness, it does so through deception rules and access doctrines, which have been identified by scholars as the backbone of American antitrust.[84]

Meanwhile, in Germany, *Leistung*, which can be translated as performance or merit, transitioned from a technical measure of force to a generalized contractual category in German law and, ultimately, to a key term of public regulation that controls and directs market logics.[85] In this way, the computational analysis offers a quantitative foundation for what intellectual historians have already suspected. While German *Wettbewerb* is semantically pulled toward performance/merit, American understandings of competition remain tethered to the agencies that regulate rates and entry. To be sure, this argument is not how much sports people actually played or watched but rather a difference in how athletic semantics spilled over into economic and legal conceptions of competition.

The doctrinal consequences of this semantic divergence are visible in at least three areas. First, regarding what constitutes abusive conduct by a dominant firm, German and EU courts have traditionally framed the inquiry in terms of *Leistungswettbewerb*.[86] The central question is whether the conduct differs from what a firm competing on the merits (that is, through superior performance rather than by disabling rivals) would do. Meanwhile, U.S. courts applying Section 2 of the Sherman Act have anchored the analysis in market structural and economic terms. They ask, for instance, whether the conduct tends to exclude an equally efficient competitor or foreclose access to an essential input. This approach reflects the institutional and access-oriented vocabulary identified in the American sub-corpus. Since EU law has progressively incorporated the equally efficient competitor test, this point is one of emphasis.

Second, the divergence on predatory pricing has historically been sharp but has narrowed in the last decade.[87] Under the *AKZO* and *Tetra Pak* line of EU cases, below-cost pricing by a dominant firm could be presumed abusive without requiring proof of likely recoupment. Since *Brooke Group* (1993), U.S. doctrine has raised the evidentiary bar considerably.[88] Plaintiffs must now show that prices are below an appropriate cost measure and that recoupment is a dangerous probability. This is consistent with a discourse tradition that treats aggressive price competition as presumptively legitimate.

---

[84] LETWIN, *supra* note 16; HOVENKAMP, *supra* note 44.

[85] VERHEYEN, *supra* note 71, at 127, 140 and 152.

[86] Helena Drewes & Tristan Rohner, *Competition on the Merits – a New Role for an Old Concept for the Analysis of Exclusionary Abuse under Article 102 TFEU*, J. COMPETITION LAW ECON. 1, 3-4 (2025); Pablo Ibáñez Colomo, *Competition on the Merits*, 61 COMMON MARK. LAW REV. 387, 378-416 (2024); Alberto Pera, *Fairness, Competition on the Merits and Article 102*, 18 EUR. COMPET. J., 229, 229-248 (2022).

[87] Brian A. Facey & Roger Ware, *Predatory Pricing in Canada, the United States and Europe: Crouching Tiger or Hidden Dragon*, 26 WORLD COMPET. 625, 625-650 (2003). For convergences in general, see RICHARD WHISH & DAVID BAILEY, COMPETITION LAW (10th ed. 2021).

[88] Brooke Group Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993).

Third, regarding access obligations and essential facilities, the EU's *Microsoft* and *IMS Health* decisions required mandatory dealing and interoperability in situations where U.S. courts, most notably in *Trinko* (2004), refused to do so.[89] Interestingly, these doctrinal differences correspond to semantic differences evident in newspaper articles.

To support this tentative interpretation, Figure 3 explicitly tracks the relative frequency of four key terms from the close readings and topic lists in both sub-corpora: *Leistung* (performance), *Sport* (sport), *Militär* (military), and *Krieg* (war). Two patterns stand out. First, the frequency of the term *Leistung* in German newspapers climbs indeed steeply from the late 1900s, peaking in the mid-1930s and remaining high into the 1940s, while the frequency of the corresponding term in U.S. newspapers stays comparatively flat. Second, sports rose in popularity on both sides of the Atlantic, but it was connected much more often to the concept of competition in Germany. In line with the literature discussed above, there was sustained growth since the First World War. References to war and the military are used to check robustness. U.S. war coverage exhibits two large waves around 1917–18 and 1941–45, with military elevated through both mobilizations. Likewise, Germany's war (*Krieg*) spikes dramatically during 1914–18, falls in the 1920s, and rises again in the 1940s. However, the corresponding peaks in the German articles are more subdued.

[89] Damien Geradin, *Limiting the Scope of Article 82 EC: What Can the EU Learn from the U.S. Supreme Court's Judgment in* Trinko *in the Wake of* Microsoft, IMS, *and* Deutsche Telekom?, 41 COMMON MARK. LAW REV. 1519, 1519-1553 (2004).

Figure 3. Relative frequency of anchor terms over time (Germany vs. United States), 1870-1945. Terms tracked: Leistung (performance), Sport (sport), Militaer (military), and Krieg (war), with corresponding English terms for the U.S. sub-corpus.

If confirmed by more in-detail legal analyses, this press-doctrine feedback loop would have significant consequences. When German lawyers and economists defend *Leistungswettbewerb*, for instance in current digital economy cases against Big Tech,[90] they introduce normative ideas of discipline and fair trial into legal analysis, the very qualities that sports made visible and that interwar newspapers had made commonplace. To this day, German lawyers (and an increasing number of non-German lawyers specializing in EU competition law) like to use a sporting analogy when introducing the key ideas behind competition law in introductory textbooks.[91] By contrast, the U.S. conversation retained a focus on institutions that enforce restraint as well as an economic vocabulary centered on deception and access rather than ethics.

If competition narratives are culturally embedded and German *Wettbewerb* and American competition do not mean quite the same thing and have not for the better part of a century, then several implications arise for contemporary practice. First, consider the intelligibility of transatlantic antitrust disputes. When EU and U.S. regulators clash over the treatment of a dominant platform, e.g. over whether refusing to share data constitutes abuse or whether a merger creates or destroys competition, it is tempting to attribute the disagreement to different economic models or political priorities. This paper suggests an additional, underappreciated source of friction: the evaluative vocabulary through which each side frames what is at stake. For example, a lawyer trained in Germany who instinctively uses the term *Leistungswettbewerb* and asks whether conduct is worthy of a fair competitor may have a different legal standard in mind than an American lawyer. Neither tradition is wrong per se, but they generate different intuitions about hard cases.

The second consequence involves algorithmic or LLM-based enforcement tools. Machine-learning screening systems that are trained to detect anti-competitive conduct and language models that are used to summarize case facts or conduct legal tests inherit the semantic assumptions of their training data.[92] For

[90] Justin Lindeboom, *Rules, Discretion, and Reasoning According to Law: A Dynamic-Positivist Perspective on* Google Shopping, 13 J. EUR. COMPET. LAW. PRA. 63, 63-74 (2022).

[91] For instance, see Bruce Lyons, *Introduction: The Transformation of Competition Policy in Europe*, *in* CASES IN EUROPEAN COMPETITION POLICY: THE ECONOMIC ANALYSIS 1, 10 (Bruce Lyons ed., 2009).

[92] In general, see the literature on biases in algorithms and LLMs, Aylin Caliskan, Joanna J. Bryson & Arvind Narayanan, *Semantics Derived Automatically from Language Corpora Contain Human-like Biases*, 356 SCIENCE 183, 183-186 (2017); Shangbin Feng, Chan Young Park, Yuhan Liu & Yulia Tsvetkov, *From Pretraining Data to Language Models to Downstream Tasks: Tracking the Trails of Political Biases Leading to Unfair NLP Models*, ARXIV (July 6, 2023), https://doi.org/10.48550/arXiv.2305.08283; Roberto Navigli, Simone Conia & Björn Ross, *Biases in Large Language Models: Origins, Inventory, and Discussion*, 15 ACM J. DATA INF. QUAL. 1, 1-21 (2023).

example, a model trained predominantly on U.S. enforcement records will have learned that harmful conduct is associated with terms such as conspiracy, foreclosure, barrier to entry, and market share. However, the same conduct described in German or EU text data may be associated with terms such as orderly competition, fair trading, and performance/merit. The risk is compounded when agencies deploy off-the-shelf commercial LLMs, i.e. models overwhelmingly trained on English-language legal text, without auditing their jurisdictional coverage. As enforcement agencies move toward algorithmic screening of conduct and LLM usage across jurisdictions, this cross-lingual semantic mismatch creates potential risks. More broadly, this paper's findings suggest an additional implication. The linguistic framing of competition accumulated over more than a century of newspaper discourse is embedded in the very training data on which modern enforcement tools are built. In short, semantic traditions rooted in the early twentieth century may quietly shape which conduct an algorithm flag as suspect today.

The final consequence is normative and necessarily speculative. A discourse in which economic rivalry is progressively rendered as national contest and martial struggle has proven vulnerable to political capture in the historical record. This does not mean that performance or sports metaphors in competition law and policy are inherently dangerous. However, legal systems that rely heavily on evaluative standards framed in athletic and/or martial terms might be particularly exposed to political intervention when geo-political circumstances shift. Today, when tariffs and AI-driven industrial policy are generating renewed pressure to frame economic competition in terms of national and strategic rivalry, the lesson is to reflect on these semantic shifts and the metaphors we use to describe genuine economic phenomena in competition law.

V. Conclusions

This paper argues that competition narratives are not a neutral backdrop to competition law. Competition narratives shape what conduct is recognized as harmful, which institutional remedies are considered natural, and how evaluative standards, such as competition on the merits or foreclosure, acquire their practical meaning. This argument has implications for interpreting historical divergences between the U.S. and Germany and for building and deploying computational tools that are central to modern antitrust enforcement. An algorithmic screening model, a text-based merger notification classifier, or a regulatory compliance assistant does not stand outside the semantic traditions documented in this paper. Rather, it encodes them.

A tool trained on a corpus saturated with one national tradition of competition discourse will reproduce the implicit frames and values of that tradition. Highlighting this issue is not a call for pessimism. Computational methods can be used to detect and measure such divergences, as demonstrated in

this paper. However, it is a reason to treat the training data of antitrust AI with the same critical attention that lawyers give to precedent. What, then, do people mean when they talk about competition? The computational analysis of approximately 1.1 million newspaper articles from the U.S. and Germany written between 1870 and 1945 yields three relevant findings.

First, semantic trajectories diverged sharply across countries and time. German competition discourse evolved from civic-commercial vocabulary in the 1870s toward organized performance metaphors during the Weimar period. Ultimately, it culminated in National Socialist rhetoric that dissolved economic categories into existential struggle narratives. Cross-lingual embeddings helped to quantify this shift. The German semantic tilt shifted from a slight war orientation to extreme militarization. In contrast, American discourse maintained a strong economic orientation, only adopting military vocabulary during wartime mobilization.

Second, sports and military vocabulary intermingled distinctly in German discourse. Dynamic topic modeling and several examples from close reading demonstrate that German newspapers increasingly linked competition to social life, championships, and performance rankings. This created semantic bridges between athletic contests and economic rivalry and produced the cultural foundation for the influential *Leistungswettbewerb* doctrine, which treats competition as a race with rules where merit is publicly recognized and consumers act as judges. By contrast, American newspapers maintained a clearer separation between economic competition (anchored to prices, firms, and regulatory agencies) and athletic competition, even as both types of vocabulary expanded.

Third, institutional vocabularies reflect deeper cultural logics. The German discourse emphasized order, fairness, and coordinated performance under public oversight. This discourse initially normalized cartel arrangements and state interventions as legitimate rule-setting rather than market distortion. In contrast, American discourse prioritized access, non-discrimination, and restraints on private power. It treated competition as a self-correcting process that required institutional protection against capture and foreclosure. These differences persisted across the period studied in this paper. Although the semantic divergences likely diminished after 1945 when the Allies helped reconstruct German competition law within the framework of a social market economy,[93] the underlying cultural logics and connotations might have prevailed. Despite decades of transatlantic cooperation and institutional convergence, the United States and Europe had regular antitrust disagreements throughout the early 21st century over merger standards, predatory pricing doctrines, and digital platform regulation.[94]

---

[93] MURACH-BRAND, *supra* note 17; Djelic, *supra* note 17.

[94] Andreas Weitbrecht, *From Freiburg to Chicago and Beyond – the First 50 Years of European Competition Law*, 29 EUR. COMPET. LAW REV. 81, 81-88 (2008); Peter Behrens, *Abschied Vom More Economic Approach?*, *in* RECHT, ORDNUNG UND WETTBEWERB 115, 115-130 (Stefan

These findings on diverging competition discourses have forward-looking implications. When politicians and regulators miscommunicate during international competition disputes, they often invoke arguments that seem universal but are actually culturally specific. These narratives embed distinctly, historically formed assumptions about what legitimate market organization looks like. In this regard, the German trajectory offers a pertinent warning. When political circumstances shift, militarized idioms of economic governance have proven vulnerable to totalitarian capture, a lesson that remains relevant at a time when geopolitical rivalry and industrial policy are again pressuring antitrust to subordinate to strategic priorities. The current AI upheaval echoes some of the historic transformations documented in this paper, such as contested control of infrastructure. A framework combining American institutional tools, such as independent agencies and access doctrines, with ordoliberal principles of competition on the merits and the special responsibility of dominant enterprises[95] may offer a more robust foundation for addressing that challenge than either tradition alone.[96] Competition narratives are not merely descriptive, they constitute the range of institutional responses that a society recognizes as legitimate when confronted with economic concentration and technological disruption. Recovering and subjecting that constitutive dimension to critical analysis is what computational methods applied to historical corpora can provide.

Bechtold, Joachim Jickeli, & Mathias Rohe eds., 2011); Rupprecht Podszun, *Digital Ecosystems, Decision-Making, Competition and Consumers – On the Value of Autonomy for Competition*, SSRN (July 17, 2019), https://www.ssrn.com/abstract=3420692.

[95] KÜSTERS, *supra* note 69, at 418–419.

[96] *See also*, in this regard Anselm Küsters & Isabel Oakes, *Taming Giants: How Ordoliberal Competition Theory Can Address Power in the Digital Age*, 141 SCHMOLLERS JAHRBUCH – J. CONTEXTUAL ECON. 149, 149-188 (2021); Manuel Wörsdörfer, *Ordoliberalism 2.0: Towards a New Regulatory Policy for the Digital Age*, 19 PHILOS. MANAG. 191, 191-215 (2020).

ANNEX A – TOPIC EVOLUTIONS FOR GERMAN AND U.S. DTM

Topic evolution for German DTM

Note: English translations are provided in square brackets.

Topic 0

1870s: fall [case], gebiet [area], berlin [Berlin], verein [association], gesetz [law], sch [shine], leistung [perfomance], bauer [farmer], Hoffnung [hope], einführung [introduction]

1880s: berlin [Berlin], fall [case], verein [association], köln [Cologne], gebiet [area], dafs [that], europa [Europe], zahlreich [numerous], linie [line], menge [amount]

1890s: verein [association], ver [ver], gesetz [law], köln [Cologne], fall [case], verhältnis [relationship], reihe [row], bedeutung [meaning], kammer [chamber], sache [issue]

1900s: kaiser [emperor], berlin [Berlin], redner [speaker], space [space], englisch [English], kreis [circle], verschieden [different], folge [consequence], stark [strong], lage [situation]

1910s: ver [ver], land [country], aufgabe [task], wort [word], konkurrenz [competition], person [person], handel [trade], maß [measure], fall [case], million [million]

1920s: form [form], erfolg [success], hoch [high], belgien [Belgium], verfügung [disposal], erheblich [significant], gebiet [area], teilnehmer [participant], stuck [piece], arbeit [work]

1930s: deutschland [Germany], kampf [battle], international [international], iſt [Iſt], deutſch [German], ver [ver], aufgabe [task], sch [shine], englisch [english], meisterschaft [championship]

1940s: deutschland [Germany], mannschaft [team], zahlreich [numerous], juli [July], zuschauer [viewers], englisch [English], stark [strong], klasse [class], angriff [attack], gauleiter [NS leader]

Topic 1

1870s: frage [ask], arbeit [work], lich [lens], verkehr [traffic], ver [ver], mark [mark], beamter [civil servant], geschäft [business], nächster [next], sinn [meaning]

1880s: frage [ask], arbeit [work], mark [mark], verkehr [traffic], ver [ver], gleich [even], nächster [next], monat [month], v [v], national [national]

1890s: frage [ask], erfolg [success], mark [mark], arbeit [work], minister [minister], kurz [short], verkehr [traffic], bekämpfung [fighting], folgend [subsequent], gewerbe [trade]

1900s: erfolg [success], linie [line], person [person], bestimmung [determination], industrie [industry], april [April], handel [trade], sache [matter], Gleich [same], versammlung [assembly]

1910s: deutsch [German], antrag [application], folgend [following], nächster [next], hoch [high], erfolg [success], verhandlung [verhandlung], folge [consequence], sache [issue], Kreis [district]

1920s: sonntag [Sunday], jahr [year], meter [meter], fahrer [driver], öffentlich [public], verein [association], handel [trade], ergebnis [result], partei [party], welt [world]

1930s: stelle [place], sport [sport], ber [ber], samstag [Saturday], sinn [sense], dezember [December], rung [tion], erſt [first], konkurrenz [competition], verhältnis [relation]

1940s: sonntag [Sunday], sport [sport], meter [meter], letzter [last], leistung [perfomance], kraft [power], minute [minute], führer [leader], treffen [meeting], stelle [position]

Topic 2

1870s: industrie [industry], letzter [last], regierung [government], englisch [English], eisenbahn [railroad], europäisch [European], gesellschaft [company], politik [policy], kaiser [emperor], werk [work]

1880s: englisch [English], kaiser [emperor], gesellschaft [company], werk [factory], industrie [industry], regierung [government], bericht [report], sür [stuff], bedeutend [meaningful], Aufgabe [task]

1890s: werk [factory], mitglied [member], gesellschaft [company], kaiser [emperor], gebiet [area], englisch [English], mai [May], geschäft [business], aussicht [prospect], Aufgabe [task]

1900s: wort [word], frage [ask], nächster [next], kraft [power], august [August], ter [ter], französisch [French], volk [people], folgend [subsequent], Zeitung [newspaper]

1910s: volk [people], werk [factory], mark [mark], berliner [Berliner], staat [state], regierung [government], name [name], könig [king], interesse [interest], gesetz [law]

1920s: kilometer [kilometer], berlin [Berlin], letzter [last], ter [ter], auftrag [order], halle [hall], international [international], sport [sport], wirtschaft [economy], president [president]

1930s: stark [strong], bild [picture], art [kind], ter [ter], düsseldorf [Düsseldorf], wett [betting], pferd [horse], ben [ben], Japan [japan], ſie [you]

1940s: ter [ter], schwer [difficult], land [country], platz [place], dresden [Dresden], gleich [even], sch [shine], fußball [football], ergebnis [result], monat [month]

Topic 3

1870s: deutsch [German], staat [state], stadt [city], kraft [power], französisch [French], partei [party], minister [minister], kammer [chamber], ausführung [execution], hiesig [local]

1880s: deutsch [German], staat [state], französisch [French], stadt [city], hiesig [local], kraft [power], london [London], absatz [paragraph], partei [party], minister [minister]

1890s: deutsch [German], hand [hand], juni [June], absatz [paragraph], ber [ber], maß [measure], kunst [art], bismarck [Bismarck], general [general], bürgerlich [bourgeois]

1900s: firma [company], grund [reason], zahl [number], deutschland [Germany], verhandlung [verhandlung], betrieb [operation], aufgabe [task], kosten [cost], Reichstag [Reichstag], fall [case]

1910s: kraft [power], gesellschaft [company], hamburg [Hamburg], berlin [Berlin], letzter [last], haus [a house], krieg [war], englisch [English], mann [man], ausstellung [exhibition]

1920s: wagen [dare], industrie [industry], stadt [city], not [need], gruppe [group], england [England], aufgabe [task], fest [firmly], köln [Cologne], stunde [hour]

1930s: sonntag [Sunday], sieg [victory], kilometer [kilometer], gebiet [area], gleich [even], england [England], einzeln [individually], verschieden [different], leben [life], runde [round]

1940s: gegner [opponent], wien [Vienna], england [England], schaft [shaft], bochum [Bochum], staat [state], mensch [person], kurz [short], schluß [finish], durchführung [execution]

Topic 4

1870s: handel [trade], mann [man], concurrenz [concurrenz], paris [Paris], wichtig [important], name [name], werth [value], früh [early], kunst [art], stelle [position]

1880s: handel [trade], letzter [last], mann [man], russisch [Russian], paris [Paris], stelle [place], besonderer [special], fremd [foreign], wichtig [important], ausländisch [foreign]

1890s: preis [price], lage [location], handel [trade], berlin [Berlin], mann [man], letzter [last], regierung [government], öffentlich [public], antrag [motion], woche [week]

1900s: verkehr [traffic], million [million], england [England], ver [ver], sch [shine], weise [way], gesellschaft [company], ſie [ſie], besonderer [important], weit [distant]

1910s: stellung [position], stadt [city], wirtschaftlich [economically], partei [party], mai [May], reich [rich], form [form], betrieb [operation], Frieden [peace], amerika [America]

1920s: deutsch [German], gewiß [certainly], deutſch [German], meldung [report], lage [location], düsseldorf [Düsseldorf], schön [nice], regierung [government], industriell [industrial], linie [line]

1930s: berlin [Berlin], vertreter [representative], jugend [youth], zahl [number], august [August], kind [child], ausland [abroad], turnier [tournament], schwierigkeit [challenge], läufer [runner]

1940s: berlin [Berlin], meisterschaft [championship], preis [price], stuttgart [Stuttgart], betrieb [operation], bann [spell], schule [school], märz [March], junge [boy], nnd [nnd]

Topic 5

1870s: million [million], leben [life], schön [nice], tarif [tariff], übrig [left over], juli [July], urtheil [judgment], reihe [row], wien [Vienna], meister [champion]

1880s: million [million], firma [company], leben [life], schön [nice], aussicht [outlook], übrig [left over], woche [week], juli [July], reihe [series], Eisen [iron]

1890s: stadt [city], partei [party], firma [company], einzeln [individually], bericht [report], leben [life], bestimmung [determination], theil [part], vorlage [template], französisch [french]

1900s: wettbewerb [competition], mitglied [member], hoch [high], werk [factory], wirtschaftlich [economically], sitzung [meeting], angebot [offer], haltung [attitude], bevölkerung [population], boden [ground]

1910s: bau [building], lich [lens], frankreich [France], weise [way], arbeiter [worker], bahn [train], iſt [Iſt], verfügung [disposal], amerikanisch [American], übrig [remaining]

1920s: wettbewerb [competition], kommend [coming], deutschland [Germany], spiel [game], mark [mark], richtung [direction], anfang [beginning], vereinigt [united], gedanke [idea], deutsche [german]

1930s: deutsch [German], preis [price], berliner [Berliner], volk [people], mark [mark], woche [week], werk [factory], industrie [industry], schweiz [Swiss], amerikanisch [American]

1940s: deutsch [German], werk [factory], gebiet [area], hamburg [Hamburg], arbeit [work], september [September], form [form], nächster [next], hoch [high], linie [line]

Topic 6

1870s: land [country], wettbewerb [competition], preis [price], reich [rich], theil [part], frankreich [France], lage [location], antrag [application], juni [June], person [person]

1880s: wettbewerb [competition], preis [price], land [country], frankreich [France], lage [location], theil [part], reich [rich], juni [June], person [person], Kreis [circle]

1890s: wettbewerb [competition], grund [reason], unlauter [unfair], weise [way], staat [state], industrie [industry], iſt [Iſt], besonderer [special], April [April], Gleich [equal]

1900s: verhältnis [relationship], letzter [last], gesetz [law], bau [building], staat [state], bedeutung [meaning], amerikanisch [American], iſt [Iſt], politisch [political], bedingung [condition]

1910s: gleich [even], verhältnis [relationship], deutschland [Germany], versammlung [assembly], industrie [industry], juni [June], lage [location], ausführung [execution], stelle [position], April [April]

1920s: grund [reason], arbeiter [worker], land [country], sieg [victory], hamburg [Hamburg], staat [state], august [August], stark [strong], tagung [conference], mann [man]

1930s: erfolg [success], arbeit [work], sieger [winner], köln [Cologne], seite [page], entscheidung [decision], sachsen [Saxony], frankfurt [Frankfurt], firma [business], Mittwoch [Wednesday]

1940s: sieg [victory], kampf [battle], köln [Cologne], münchen [Munich], leben [life], tor [goal], rahmen [frame], ungarn [Hungary], entscheidung [decision], bedeutung [importance]

Topic 7

1870s: interesse [interest], hoch [high], folge [consequence], einzeln [individually], hand [hand], art [kind], volk [people], auge [eye], öffentlich [public], folgend [subsequent]

1880s: hoch [high], folge [consequence], interesse [interest], einzeln [individually], hand [hand], fabrik [factory], fürst [prince], auge [eye], folgend [subsequent], industriell [industrial]

1890s: arbeiter [worker], reichstag [Reichstag], interesse [interest], lich [lens], hoch [high], juli [July], person [person], schutz [protection], reich [Empire], ausstellung [exhibition]

1900s: deutsch [German], neu [new], köln [Cologne], mark [mark], regierung [government], interesse [interest], seite [page], partei [party], vertreter [representative], verband [association]

1910s: frage [ask], preis [price], grund [reason], flug [flight], bedeutung [meaning], verein [association], space [space], international [international], geschäft [business], rung [rung]

1920s: januar [January], antrag [application], oktober [October], politik [policy], gefahr [danger], gleich [even], ver [ver], möglichkeit [possibility], werk [work], märz [March]

1930s: platz [place], hoch [high], meister [master], besonderer [special], juli [July], grund [reason], folgend [following], ten [ten], lage [situation], straße [street]

1940s: welt [world], sachsen [Saxony], titel [title], besonderer [special], kreis [circle], grund [reason], folgend [following], begegnung [encounter], mädel [girl], wirtschaft [economy]

Topic 8

1870s: arbeiter [worker], seite [page], england [England], nation [nation], rußland [Russia], erfolg [success], richtung [direction], verwaltung [administration], zahl [number], allgemein [general]

1880s: england [England], arbeiter [worker], art [kind], zahl [number], seite [page], gering [small amount], markt [market], lang [long], nation [nation], erfolg [success]

1890s: nächster [next], stelle [place], zahl [number], art [kind], lang [long], ausführung [execution], preußisch [Prussian], wichtig [important], haus [house], bau [construction]

1900s: woche [week], leipzig [Leipzig], land [country], geschäft [business], entwurf [draft], leben [life], voll [full], wert [value], architekt [architect], welt [world]

1910s: wert [value], firma [company], unlauter [unfair], gebiet [area], zukunft [future], öffentlich [public], besonderer [special], kammer [chamber], stark [strong], architekt [architect]

1920s: geschäft [business], v [V], absatz [paragraph], haus [a house], men [men], folgend [following], million [million], derartig [such], juli [July], woche [week]

1930s: meter [meter], land [country], spiel [game], leistung [perfomance], führer [leader], reich [rich], stadt [city], frage [ask], teilnehmer [participant], verein [club]

1940s: spiel [game], gau [gau], meister [master], erfolg [success], bereich [area], januar [January], volk [people], teilnehmer [participant], runde [round], Leipzig [Leipzig]

Topic 9

1870s: deutschland [Germany], ausstellung [exhibition], bahn [train], weise [way], verschieden [different], frei [free], grund [reason], neu [new], stellung [position], wort [word]

1880s: deutschland [Germany], ausstellung [exhibition], verschieden [different], berliner [Berliner], wort [word], weise [way], zweck [purpose], stellung [position], politisch [political], bahn [lane]

1890s: zweck [purpose], million [million], berliner [Berliner], verschieden [different], land [country], deutschland [Germany], wort [word], stellung [position], sitzung [meeting], versammlung [assembly]

1900s: hiesig [local], preis [price], verein [association], unlauter [unfair], arbeit [work], stadt [city], gebiet [area], heutig [today], jahr [year], billig [cheap]

1910s: wettbewerb [competition], england [England], seite [page], flieger [plane], juli [July], neu [new], frei [free], art [kind], monat [month], welt [world]

1920s: fall [case], preis [price], prozent [percent], frage [ask], leben [life], lang [long], nächster [next], wert [value], bahn [railway], wort [word]

1930s: letzter [last], punkt [point], mann [man], gruppe [group], minute [minute], frankreich [France], strecke [route], hand [hand], höhe [height], freude [joy]

1940s: mann [man], wettbewerb [competition], punkt [point], jugend [youth], krieg [war], gruppe [group], reich [rich], aufgabe [task], oktober [October], waffe [weapon]

Topic evolution for U.S. DTM

Topic 0

1870s: business, trade, place, government, labor, bill, oil, world, fine, thing

1880s: labor, people, business, trade, bill, government, general, system, place, thing

1890s: road, government, country, street, party, high, land, duty, question, hand

1900s: american, work, land, committee, ing, life, mile, money, month, car

1910s: power, free, land, mile, street, club, party, result, america, order

1920s: ing, month, sale, hand, local, thing, round, john, coal, result

1930s: washington, event, champion, ball, week, place, club, district, boy, john

1940s: club, season, champion, long, production, south, red, bob, frank, trophy

Topic 1

1870s: price, state, people, mile, cost, free, pacific, success, iii, necessary

1880s: city, country, free, land, article, wage, product, pacific, cost, union

1890s: company, silver, wage, value, cost, union, demand, protection, office, capital

1900s: government, product, people, case, high, woman, department, match, market, order

1910s: government, competition, money, woman, cup, average, fine, chief, dealer, democratic

1920s: united, week, large, prize, ment, match, american, city, business, ball

1930s: competition, man, game, large, city, class, demand, competitor, financial, tive

1940s: united, woman, ball, district, title, line, foot, department, play, golf

Topic 2

1870s: country, market, stock, class, land, course, school, england, home, farmer

1880s: man, state, company, market, class, home, england, necessary, woman, south

1890s: competition, race, de-, england, woman, yard, department, mill, open, strong

1900s: place, club, race, labor, competition, national, general, pro-, hour, judge

1910s: present, fact, tho, united, president, con-, people, race, tariff, class

1920s: public, foreign, yard, night, house, money, final, british, court, conference

1930s: field, second, com, school, contest, people, money, case, private, tonight

1940s: team, field, com, round, people, entry, board, sport, month, present

Topic 3

1870s: competition, american, street, low, states, question, prize, duty, team, open

1880s: competition, american, public, road, street, low, duty, report, national, hour

1890s: president, interest, public, free, course, service, washington, association, report, john

1900s: viii, home, low, president, present, washington, duty, water, tho, service

1910s: prize, committee, game, fair, price, high, bill, contest, law, general

1920s: race, fact, course, low, tion, government, war, class, bill, fair

1930s: government, ton, match, national, player, night, course, law, service, company

1940s: national, man, service, business, government, company, committee, army, course, right

Topic 4

1870s: public, tor, ton, week, party, board, foreign, gold, yard, purpose

1880s: party, stock, week, committee, foreign, point, right, question, chicago, result

1890s: price, city, ton, stock, world, result, purpose, night, iron, necessary

1900s: week, trust, trade, interest, line, question, de-, san, standard, night

1910s: commission, thing, house, home, sale, question, white, life, low, purpose

1920s: competition, team, people, country, school, washington, en-, marine, development, distance

1930s: president, work, world, court, commission, de-, mile, college, hand, number

1940s: competition, second, city, place, final, work, little, class, end, park

Topic 5

1870s: high, viii, point, little, member, matter, kind, tue, bid, term

1880s: rate, ing, ton, high, little, matter, mile, protection, member, reason

1890s: united, rate, law, house, mile, second, trust, case, member, condition

1900s: man, company, railroad, capital, free, short, port, agent, re-, building

1910s: work, month, little, trade, ton, market, states, product, department, case

1920s: state, line, president, ton, industry, com, trade, law, street, railroad

1930s: country, point, tournament, low, winner, league, season, play, interest, wage

1940s: war, league, ing, state, week, point, country, ton, member, prize

Topic 6

1870s: aud, city, ana, road, rate, sale, fact, power, order, charge

1880s: price, aud, fact, order, sale, ana, power, industry, profit, prize

1890s: american, committee, trade, order, point, thing, people, pacific, large, prize

1900s: large, business, united, rate, point, ton, association, public, stock, world

1910s: large, team, washington, corporation, steel, board, interest, ship, profit, athletic

1920s: man, committee, high, national, star, meeting, afternoon, real, charge, short

1930s: business, star, line, foot, little, federal, fact, general, nation, railroad

1940s: game, world, program, price, nation, college, association, air, navy, fact

Topic 7

1870s: tho, railroad, house, line, money, number, com, second, contract, value

1880s: tho, house, railroad, line, money, com, tion, number, month, life

1890s: money, states, place, power, home, little, work, month, fact, board

1900s: state, oil, foreign, little, states, street, record, town, coal, game

1910s: ing, field, city, american, stock, com, public, county, record, number

1920s: tho, point, place, club, number, condition, commission, power, car, champion

1930s: ing, high, final, championship, industry, record, woman, association, miss, victory

1940s: high, washington, night, boy, george, john, industry, home, miss, job

Topic 8

1870s: man, large, work, interest, present, law, united, ing, general, committee

1880s: large, law, work, tariff, present, interest, united, county, case, ment

1890s: man, state, line, ing, week, present, market, labor, low, general

1900s: law, condition, school, tion, course, class, team, district, steel, john

1910s: tion, week, man, course, system, school, world, national, point, george

1920s: states, rate, price, company, home, labor, woman, de-, george, market

1930s: price, american, state, race, code, labor, round, home, power, rate

1940s: american, star, record, race, states, large, victory, tion, house, university

Topic 9

1870s: company, hand, president, subject, capital, national, evening, de-, iron, bond

1880s: president, hand, capital, subject, court, de-, evening, con, room, property

1890s: bill, class, product, business, ment, railroad, number, sale, county, school

1900s: tariff, road, city, country, price, county, number, com, bill, building

1910s: state, company, business, country, war, line, second, rate, railroad, condition

1920s: cost, second, tariff, cup, mile, open, foot, land, member, service

1930s: team, united, states, member, trade, cup, public, open, board, yard

1940s: school, event, tournament, winner, president, player, open, bill, championship, yard