

April 2026

Computing Diachronic and Synchronic Semantic Drift: The Case of *Redpill* on Reddit, 2012–2024

**An NLP-Based Framework for Socio-Semantic
Network Analysis of Digitally Native Language &
Semiotics**

Moritz Schwenger

Master's Thesis supervised by Sylvain Parasie

Second Member of the Jury: Jean-Philippe Cointet

Master in European Affairs

Digital New Technology and Public Policy

Preface

Why you should read this paper

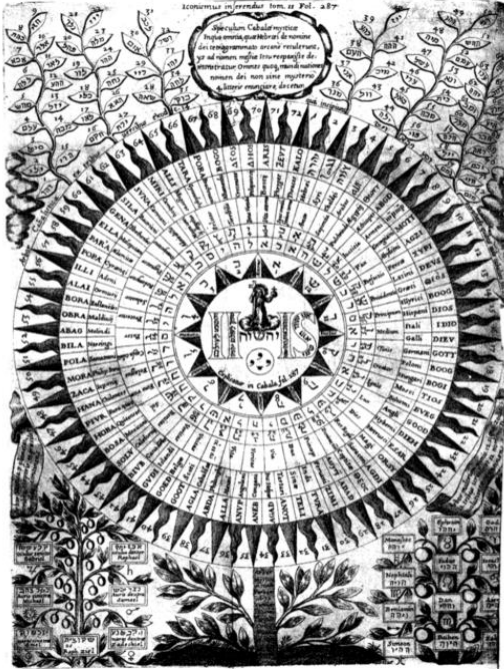


Figure 1

“Language bears within itself the necessity of its own critique.” **Jacques Derrida**

To the Postmodern reader the unstable volatility of all meaning should not only be a philosophical constant but a source of enjoyment. Among the many examples that come to mind when contemplating this moment in time are the wanton mischief of Thomas Pynchon's pseudo-encyclopedic descent into madness in *Gravity's Rainbow* (1973), and the metafictional demise of the poet John Shade in Nabokov's *Pale Fire* (1962), both of which revel in playful ambiguity, and the deferral of interpretation. From its literary to its theoretical manifestations, the postmodern movement has produced a vast body of work, probing the dangers of Enlightenment sincerity and the hierarchical organization of meaning.

Like many works of academic writing, the following thesis finds its inspiration equally in the pop-cultural anxieties of its time, as it does in the joy of betrayal of what came before. The anxiety

Diagram of the Names of God (17th century), engraving by Athanasius Kircher. Public domain image from Wikimedia Commons, originally published in *Oedipus Aegyptiacus* by Athanasius Kircher. Wikimedia Commons, file: *Kircher Diagram of the names of God.png*. Public domain.

of our time, or so it would seem, has taken language itself as its primary object.

As early as the publication of Ludwig Wittgenstein's *Tractatus Logico-Philosophicus* in 1922 Weimar Germany, and amidst the implosion of his world, of the Viennese continental system and the final aftershocks of Europe's *fin-de-siècle* hegemony, language emerged as a central site of philosophical crisis. The converging forces of the latter half of the twentieth century have since intensified this condition, producing an unprecedented technological and analytic fixation on said subject matter, one that posits language simultaneously as the medium, the object, and the limit of contemporary epistemological inquiry.

Now, almost three decades after the movement's peak, the academic tradition is reassessing the political, cultural, and social consequences of the last century's second half. Many of which include post-truth mass media, the (ab)uses of publication technologies, and the poly-centric world order. Ironically postmodernism, which has

centered its critique around the shortcomings of language, has risen concurrently to perhaps the most text-centered and semiotically charged space in human history, the Internet.

Therefore, this thesis wants to trace the impact of new linguistic patterns, resulting from this new public sphere and its impact on liberal societies. To contemplate cyberspace as a technological substrate, what Gilles Deleuze described as product of a generalized crisis in relation to all environments of enclosure, is no longer the premeditated eschatology of late postmodern thinkers, but rather the basis of our lived experiences. This fragmented, plurocentric world has been profoundly shaped by these technologies, which have in turn contributed to the emergence of a digitally mediated public sphere (Berthon, 1998). Other thinkers, such as René Girard with his concept of mimetic desire, have likewise left a lasting imprint on the sociotechnological imaginaries of the early Internet. Yet the central fallacy of its primordial days lay in the belief that it constituted a natural extension of the bourgeois public sphere, of sustaining rational-critical debate, or and consensus formation through deliberative exchange among autonomous subjects.

By examining processes of new linguistic forms, this study demonstrates not only how humans shape language through use, but also how technological systems (content moderation, algorithmic curation) actively shape linguistic adaptation and innovation. Still, the thesis remains attentive to the epistemological limitations that diachronic semantic drift may impose on the future development of language- and speech-dependent technologies, including the Internet itself. This technological short circuit might like no other necessitate to not only advance technological tools but recenter the importance of organizational structures to progress.

One need not look far into the history of knowledge to encounter earlier

experiments in systematic knowledge organization as well as the tension between machines and human cognition, ergo language. Kircher's Renaissance volvelle, *Diagram of the Names of God* (Fig. 1) operates as a mechanism for the externalization of cognition, materializing thought as a manipulable combinatorial system. The device presupposes that language is not an arbitrary medium but a privileged site in which the structure of reality becomes expressible.

Contemporary AI systems and the web inherit the formal ambition of universal interconnection while simultaneously evacuating its metaphysical grounding, relying vastly on the quality of human language, reorganized within this technological architecture. Where Humans once searched attunement to the dynamics of meaning within a presumed order of being, this thesis remains weary of a general degradation of language's epistemological, and aesthetic functionality.

The *Diagram of the Names of God* finds a striking, albeit more tongue-in-cheek, echo in Arthur C. Clarke's short story *The Nine Billion Names of God*. In Clarke's tale, two American computer scientists undertake the arduous journey to the Dalai Lama's remote Tibetan monastery to assist the monks in completing the monumental task of enumerating every name of God or knowledge itself. What would otherwise require centuries to achieve is now accomplished in a matter of days, through the computational power of the technologically adept visitors. Yet the Western skeptics remain anxious, uncertain of the consequences should nothing occur upon completion. As they descend back to the center of their world, one of rational machines, the narrative reaches its enigmatic climax: *the stars begin to disappear*.

Table of Contents

Preface.....	1
Why you should read this paper	1
Table of Contents.....	3
Acknowledgment.....	5
Citation and Source Conventions	5
AI Disclosure	5
Introduction.....	6
Existing Studies.....	12
Original Contributions.....	14
Theoretical Framework.....	15
A History of Semantic Drift	16
State of the Art.....	17
Qualitative Methods.....	18
Quantitative Methods	20
Human-Machine Language Coevolution and Constraint Shaped Registers	23
Technologically Dependent Semantic Drift.....	23
The Written Speech Paradox.....	23
Semiotics of the Cyberspace	25
Mainstreaming Process.....	31
Critical Discourse Analysis and Netnographic Approaches	32
Features of Coevolved Language.....	34
Neologisms.....	34
How to Adapt Historic Methods.....	36
Main Part.....	37
Case Study Setup	38
Taking the Redpill.....	38
Criteria for a Suitable Word Token	38
Why Redpilled and Its Variations?	40
Why Reddit?	42
Data Collection Corpus Building	44
Methodology and Preliminary Results.....	47
Semantic Drift Analysis.....	47
Nearest-Neighbor Extraction.....	51
Dictionary	51
Semantic Field Classification.....	54
Results	58
Proximity of concept words over time periods.....	58
Proximity of concept words over Subreddits.....	62
Correlation Matrix.....	69
Limitations	71
Final Results	72
Lexicographic Profile and Semantic Field Entry.....	72
Discussion	77
Lexicographic History on Reddit.....	77
Matrix-original use.....	79
Manosphere use.....	80
#Gamergate: Politicizing the Redpill.....	82
Donald Trump & Redpilling.....	84

Manosphere Terrorism.....	88
Counter-Discourse, Virality, and Open Signifier	90
Recommendations	92
<i>References</i>	<i>94</i>
Sources	95
Appendix.....	109
Data Set Audit	110
Abstract.....	117

Acknowledgment

A special thank you to Sylvain Parasie of Sciences Po Paris for supervising this thesis, and to Jean-Philippe Cointet of Sciences Po Paris for serving as co-juror. I am grateful for their input and suggestions which have informed the entire project. I would also like to thank Stéphane Baele (UCLouvain) and Lewys Brace (University of Exeter) for their guidance on the coding framework and methodological set-up. Thanks also to Stephen Paul Atkinson of the École Normale Supérieure Paris and Duke University for editing and reviewing the first complete draft of this research project. Finally, I am indebted to the Zotero team at Sciences Po and the librarians at Saint Thomas for confirming and cross-checking multiple sources throughout the research process.

Citation and Source Conventions

All sources are hyperlinked throughout the text and accessible through these links wherever cited. Direct quotations are placed within double quotation marks; paraphrased material is indicated by proximate citation in the format (Author, Year). The origin and extraction pipeline of the dataset is disclosed within the text and can be made available by contacting the author. Citations drawn from the study's own dataset are explicitly marked as such. The dataset was analyzed in Google Colab, and all graphs were produced in Python within the same environment. Additional graphs and a 3D vector map not featured in the body of the thesis can be found in the Appendix or accessed under this [link](#)

AI Disclosure

No generative AI was used for text production. Spell-checking and editorial suggestions were drawn from Microsoft Word's built-in spell checker, Grammarly, and after completion verified by Claude AI via the prompt (“Without modifying the text, return a list of all spelling and grammatical mistakes”, identifying 22 unseen mistakes). Within Google Colab, the integrated Gemini AI was used throughout the coding process whenever errors occurred, which made it possible to build a stable and error-free pipeline. AI suggestions were accepted sparsely whenever the manual code did not run as intended and were applied on demand by the system; for this reason, individual AI interventions are not exhaustively documented. Graphs were produced in Python on Google Colab, with occasional Gemini-assisted corrections to identify the appropriate scale or formatting. All comments were stored within an HTML file that allows the corpus to be filtered and searched by keyword, supporting the identification of broader trends within the dataset. Notebook LM was used sparingly during the literature review to reconfirm prior readings of sources (i.e., to verify that earlier interpretations of a given text remained accurate). Due to word-count limitations, not all comments analyzed in the qualitative section could be reproduced in the body of the thesis; these are available on request via moritz.schwenger@sciencespo.fr

Introduction

In May 2020, the technology magnate Elon Musk posted a three-word imperative to his then 34 million Twitter followers: "[Take the red pill](#)." Within minutes, Ivanka Trump, daughter and senior advisor to the sitting President, replied: "[Taken](#)." Yet the term at its center had already undergone a transformation so extensive that no single reading of the injunction could be authoritative.

What had originated as an epistemological metaphor out of the Wachowskis' *The Matrix* (1999), in which the hacker Thomas Anderson (Keanu Reeves), operating under the alias Neo, is offered a choice between comfortable ignorance and the unbearable apprehension of the real, had, by that point, traversed a remarkable discursive itinerary ([Taylor for BBC](#), 2021). In large numbers, the term had first been appropriated as the ideological totem of the *manosphere*, consolidated around the subreddit *r/TheRedPill* and its pseudonymous founder *pk_atheist*, later identified as Robert Fisher, a New Hampshire state legislator whose clandestine advocacy of anti-feminist sexual strategy would itself become an object of journalistic exposé. It had subsequently migrated into the lexicon of the incel community, where its deployment oscillated between ironic detachment and the endorsement of fatalistic, at times violent, worldviews. Later being transposed into conspiratorial registers where *redpilling* denoted not the acquisition of gendered knowledge but a millenarian awakening to hidden structures of power. Now as these communities have mostly been shut down the term is predominantly used in counter discursive spheres or favorably replaced by its successor *blackpill* ([Tiffany](#), 2021). By the time of Musk's tweet, *redpill* had however accrued so many competing significations that it had become, in effect, semantically overburdened. "*The Alt-Right has lost control of Redpill*", notes American journalist Kaitlyn Tiffany a year after Musk's infamous tweet in *The Atlantic*.

The centripetal migration of peripheral online communities toward the discursive mainstream constitutes one of the more extensively documented phenomena in contemporary digital sociology ([Ribeiro et al.](#), 2021); ([Nagle](#), 2017); ([Munn](#), 2019). Over the past decade, scholars and policymakers alike have attended to the accelerating and unprecedented transposition of fringe political ideologies into the center of the public sphere. As these communities migrate, they equally assimilate into the communicative ecologies they enter, ipso facto transforming both themselves and the discursive environments that receive them.

Inevitably, as the Internet remains a fundamentally text-centered technology, these movements are propelled and their formations mediated by the language that surrounds them. Their shared signs and metaphors, their artefacts of collective sense-making, their codes of mutual recognition, their irony and humor all travel with them ([Androutsopoulos](#), 2011).

Equally, within the tradition of the bourgeois public sphere (Cf. Jürgen Habermas), language and taste had been ascribed a particular role in this dynamic, one through which the construction of the self and the communities that self might inhabit become expressed (Bourdieu, 1991). The history of publication media, the expressive norms of democratic discourse, and the aesthetics of the sayable in good taste have been one of the major imaginaries of the liberal tradition, through which its inherent tensions of constrained freedoms, not least the freedom of speech, partly become mediated. Adapting to the technological affordances of their times, the fundamentally textual Internet, thus the digital public sphere ([Habermas, 2022](#)) prompted the radical democratization of publication media and a disembodied mode of discourse, while necessitating the emergence of regulatory tools (*see* Features of Coevolved Language). To cash in the insight already documented, such as the migratory patterns as well as the culture of these online communities, one ought to return to the underpinning structures of this change, language itself.

Among the most distinctive features of this new communicative environment is its capacity to generate entirely novel lexical forms ([Würschinger & McGillivray, 2024](#)), such as *redpilled*. These digitally native neologisms, terms born within and contingent upon platform-specific conditions, emerge at a pace without precedent in the history of natural language ([Crystal, 2001](#)); ([Herring & Androutsopoulos, 2015](#)). Unlike their analogue antecedents, whose semantic trajectories unfold across long time spans, these terms compress the full cycle of coinage, diffusion, reappropriation, hence semantic transformation into quick time frames. Still, when these terms may cross community boundaries, which they often do along new literary forms such as memes, their meanings also adapt to local use. They might be ironized, (mis)appropriated, ideologically recharged, or simply become meaningless. Simultaneously and under conditions such as the Internet's semiosphere ([Bankov, 2022](#)) neither speaker nor interlocutor can be certain of the deployment register. Therefore, this process, which traditional linguistics calls semantic drift, occurs at unprecedented scale in digital environments, and with political consequences to which the following study will attend in unmatched scale and scope.

Despite the breadth of existing research, several lacunae persist. No diachronic study has traced the full migratory trajectory of a single token across an ideologically diverse ecosystem. Studies have either examined community language synchronically ([McGlashan & Krendel, 2024](#)) or tracked user migration without measuring what happens to the words those users carry ([Ribeiro et al., 2021](#)). The technical feasibility of diachronic distributional semantics has been established ([Hamilton et al., 2016](#)); ([Würschinger & McGillivray, 2024](#))

yet never applied to ideologically charged digitally native tokens. *Redpilled* itself has been studied almost exclusively through qualitative methods, with the OILab frequency analysis ([Hagen, Tuters & Wilson](#), 2020, not peer-reviewed) constituting the only prior quantitative attempt and applying no distributional semantic modeling. No existing framework connects semantic drift to ideological function in enabling cross-community migration. Though having been described qualitatively ([Phillips & Milner](#), 2017), it has never been integrated into a computational diachronic framework (Cf. Existing Studies for full overview). These gaps bear on questions beyond linguistics. If ideologically loaded language can mainstream while remaining opaque to moderating bodies, this carries implications for content governance and democratic discourse. If content moderation tools become the new arbiters of the sayable, the registers of constrained expression they produce raise their own moderative, aesthetic and democratic questions.

This thesis pursues two aims. First, it provides an unprecedented mixed-method approach that captures both the diachronic and synchronic dimensions of semantic change within a single digitally native lexical item. It introduces a framework for tracing the diachronic migration of a speech token across a reconstructed subreddit ecosystem, combining Word2Vec distributional semantics with Procrustes alignment to measure semantic drift over time, while simultaneously attending to the synchronic variation of the term's meaning across ideologically distinct communities at any given temporal cross-section. Second, it applies this framework empirically to *redpilled* across 25 subreddits and approximately 966,000 comments spanning from 2012 to December 2024.

Henceforth, the thesis argues that as *redpilled* migrates from its origin communities into mainstream and counter-discourse spaces, a decoupling occurs between its ideological potency, i.e the capacity to convey a specific political program, and its semantic charge, i.e the capacity to trigger recognition and affective response. The term might become culturally legible but ideologically plural. It can simultaneously function as sincere commitment, ironic mockery, and content-neutral slang. This decoupling is not accidental but is produced by specific structural features of digital communication, among them the absence of paralinguistic cues, the affordances of anonymity, and the acceleration effects of algorithmic curation. The paper partly attends to the influence that external moderation events may exert on the semantic stability of a term. When origin communities are banned or quarantined, the deliberative spaces in which a token's meaning was negotiated are dissolved. The thesis therefore asks whether the suppression of these communities, which could once argue for the authoritative meaning, transforms *redpilled* into a floating signifier into which various

ideological projects and sensibilities may be projected, mythologizing the term in a manner that is no longer anchored to any single community's discursive authority.

H1: Decoupling of potency and charge
As <i>redpill</i> migrates from origin communities (<i>r/TheRedPill</i> , <i>r/Incels</i>) toward mainstream and counter-discourse spaces (<i>r/AskReddit</i> , <i>r/IncelTears</i>), its ideological potency becomes pluralistic while its semantic recognizability and affective charge increase, thus producing a term that is culturally legible but ideologically inert in mainstream contexts.
H2: Platform moderation as a driver of semantic acceleration
Major content moderation events (quarantine of <i>r/TheRedPill</i> in 2018, ban of <i>r/Incels</i> in 2017, ban of <i>r/Braincels</i> in 2019) function as drivers of semantic drift (or reorientation), forcing the term into new communities and accelerating its transformation.

To apprehend these dynamics, the thesis proposes several conceptual tools developed at length in the theoretical framework. Strategic semantic drift designates semantic change shaped by platform architectures and exploited for political purposes. Constraint-shaped registers (CSR) denote the distinctive linguistic forms that emerge as creative adaptations to technical constraints such as content moderation and algorithmic filtering, subsuming what the literature has variously called computer-mediated discourse, netspeak, and algospeak. The Written Speech Paradox describes digital text that adopts the immediacy of spoken conversation without its paralinguistic anchors, producing structural conditions for semantic instability. The distinction between ideological potency and semantic charge differentiates a term's capacity to convey a specific political program from its capacity to elicit recognition and affect, two dimensions this thesis contends can decouple as a token migrates, producing what the study calls a *plurosemantic political vessel*. Finally, a token such as *redpilled*, tethered to no analogue referent, is susceptible to what [Eco](#) (1976) termed unlimited semiosis. Each reappropriation, such as counter-discursive or ironic inversion, adds further strata of meaning without fully displacing those that came before.

Key Definitions
Digitally native neologism (DNN)
lexical items that emerge from and are contingent upon digital communication environments; they carry meanings hardly intelligible outside their digital context of origin. Distinguished from digitally <i>modified</i> analogue language (e.g., "cloud," "feed") and from semiographic communication (emojis, memes).

Semantic drift (diachronic vs. synchronic)
Diachronic: the evolution of a word's meaning over time. Synchronic: the coexistence of different meanings across communities at a single point in time. This thesis combines both axes.
The manosphere
a networked ecosystem of online communities organized around anti-feminist and masculinist ideologies, including Men's Rights Activists (MRA), Pick-Up Artists (PUA), Men Going Their Own Way (MGTOW), and involuntary celibates (incels). Characterized as a dynamic migratory ecosystem with fading barriers between moderate and extreme subgroups.
Incel (involuntary celibate)
a self-identified community organized around the belief in biologically determined romantic/sexual exclusion. Distinguished by its fatalistic "blackpill" ideology and a documented corpus of community-specific jargon.
Content moderation event
a platform governance intervention (quarantine, ban, removal) that disrupts a community's communicative infrastructure and forces user and linguistic dispersal. Treated in this thesis as a potential causal inflection point in semantic drift.
Redpilled
brief working definition here (expanded in the full lexicographic profile). A digitally native ideological token originating from <i>The Matrix</i> (1999), reappropriated by masculinist communities to denote ideological awakening, and subsequently migrating across the political spectrum.

The study confines itself to Reddit, from the first attestation of *redpilled* in *r/TheRedPill* in 2012 through December 2024, drawing on the Pushshift.io archive for its longitudinal depth. The linguistic scope is restricted to English-language comments containing *redpill* and its morphological variants, all normalized to a single token for distributional modeling. The study does not track the term across adjacent platforms such as 4chan, Twitter/X, or TikTok, nor does it undertake sentiment analysis or toxicity classification. It makes no causal claims about radicalization pathways, only about how the term's semantic profile transforms as it migrates across communities. A constitutive limitation must be acknowledged. Approximately 182,000 entries, concentrated within *r/TheRedPill*, *r/Incels*, and *r/Braincels*

during their peak and pre-ban periods, were removed prior to archival capture by Reddit's moderation infrastructure. This loss is intrinsic to the dataset and not correctable post hoc.

The thesis proceeds by surveying the existing literature on digital linguistics, manosphere discourse, and computational approaches to semantic change, delineating the specific research gaps this study addresses. It then establishes the theoretical framework within which the subsequent analysis is situated, tracing a path from the tradition of diachronic semantics through the distinctive conditions of digitally mediated language /constraint-shaped registers, semiospheric architecture *et cetera*). Part III turns to the case study design, furnishing the criteria by which a suitable token was identified, the justification for selecting *redpilled* and its morphological family, and the rationale for Reddit as the corpus environment from which the term's migratory trajectory can be reconstructed. Subsequently, it sets out the methodological apparatus, detailing the data collection pipeline and the Word2Vec with Procrustes alignment framework through which semantic drift is operationalized and measured. Arriving at presenting the empirical findings and interleaves them with a qualitatively sustained semantic history of the term, drawing on illustrative comments from the dataset to anchor the computational results in the discursive realities from which they are abstracted. Furthermore, it then returns to the theoretical framework to evaluate the two hypotheses considering the evidence adduced, attending to both their confirmations and their limitations. The thesis concludes by considering the implications of these findings for content governance, for the epistemological integrity of text-centered technologies, and for the broader question of how democratic regimes might attend to a public sphere whose dominant medium of expression is perpetually, and perhaps constitutively, in accelerating flux.

Existing Studies

Both quantitative and qualitative approaches have been undertaken to document the linguistic changes concurrently occurring on digital platforms. The technical feasibility of diachronic corpus analysis has been performed in the following precedents: In their 2024 study [Würschinger and McGillivray](#) trace the semantic change and socio-semantic variation of COVID related neologism with a Word2Vec Procrustes alignment approach, serving as the closest methodological parallel to this study. Equally, the diachronic intensification of violent language across the incel ecosystem has been observed by [Baele, Brace & Ging's](#) 2023/4 cross platform study via a predefined dictionary approach. This thesis has particularly been made possible through Brace personal help to establish the subsequent methodology. While [Karjus and Cuskley](#) (2024) index the linguistic divergence across polarized social media

communities, showcasing the viability of synchronic cross-community semantic divergence mapping.¹

The mansphere and incel ecosystem, which ideologically married various fringes such as MRA (Men's Rights Activism), PUAs (Pick Up Artistis), and MGTOW (Men going their own way), has been ascribed radical tendencies with the communities' sense-making circulating around a shared anti-feminist framework ([Ging, 2019](#)). This dynamic migratory ecosystem ([Ribeiro et al, 2021](#)), has been documented in large-scale empirical studies, reinforcing the idea of fading barriers between modern and extreme subgroups on Reddit. A shared transgressive memetic aesthetics has been ascribed to these Subreddits, often tracing their cultural-historical genealogy to early 4chan and other Far Right subcultures ([Nagle, 2017](#)), to which, after content moderation decisions, they have partly returned or migrated elsewhere.

The community's language has been ascribed an increasing hostility and violence, rewarding gender-based and misogynistic violence ([Farrell et al., 2019](#)). Specific jargon (see Digitally Native Neologism) has emerged as an in-group marker and synchronic linguistic feature, while complement keywords have been attributed as distinctive to individual mansphere communities with a corpus-linguistic keyword analysis ([McGlashan & Krendel, 2023](#)). Both the baseline features of misogynistic discourse in incel communities ([Jaki et al., 2019](#)), the on hierarchical group dynamics of linguistic boundary policing ([Solopova et al., 2025](#)), and the community's relationality to gender and race ideologies ([Heritage, 2023](#))² are studied, leading to the most comprehensive incel lexicon to date ([Gothard et al., 2021](#)). The Incel and mansphere communities are understood as constantly evolving ideological communities that constantly renegotiate their focuses. Generally, they feature traditionalist views on gender and promoting anti-feminist viewpoints. At times they mingle and integrate racist, white nationalist, queerphobia, and conspiratorial discourses. IT has become emblematic of the contemporary crisis of male isolation. The aforementioned first diachronic cross-platform and corpus-based study of the incelosphere's extremist language by [Baele, Brace & Ging](#), while working with a predefined dictionary approach, remains the most pertinent predecessor to this study, which will apply it to one single lexical item. A subsequent intervention by [Brace, Baele and Ging \(2024\)](#) extended the analytical purchase of this framework to the question of mixed, unclear, and unstable ideologies within the incelosphere.

¹ Note that while these studies demonstrate the technical feasibility of the approach, none applies it to ideologically charged fringe-to-mainstream language migration.

² From abstract; not accessible via Sciences Po Library

The most proximate study with respect to the specific lexical item under examination is that of [Hagen, Tuters and Wilson \(2020\)](#), whose [OILab](#) analysis constitutes the only prior attempt to track redpill as a discrete token across Reddit subreddits using Pushshift data. Their localization of the term until 2020 is similarly mirrored, not congruent, in the current data set. The not peer-reviewed study does not, however, apply any form of distributional semantic modelling, leaving the question of the term's semantic profile and local meanings unaddressed. Qualitative studies have tried to address this gap. [Waśniewska \(2020\)](#) offers a complementary linguistic perspective, analyzing the cultural symbolism and conceptual metaphor embedded in the pill-family terminology within incel discourse, attending to the ideological work performed by the *redpill* metaphor at the level of the individual sign. Incel neologisms function as cryptolectal boundary markers, producing in-group legibility while maintaining opacity to outsider. Furthermore, research has established *redpill* as a place of sense making with a sourceful cultural history ([Bratich, 2023](#)), also for women ([Hoebanx, 2024; here 2025](#)). The term has also been granted an important role as epistemological metaphor in conspiratorial discourse ([Van Veeren, 2025](#), not yet published but available via Academia.edu), going beyond its association with gender politics. Find other note-worthy studies under sources³.

Original Contributions

1. First diachronic Word2Vec + Procrustes alignment study of a manosphere token	Farrell et al. (2020) applied word embeddings to manosphere jargon synchronically only. Hamilton et al. (2016) established the diachronic method but on historical literary corpora, not digitally native ideological tokens.
2. Combines diachronic and synchronic distributional axes in one framework	Karjus et al. (2024) used a comparable dual-axis approach but on two communities over two timepoints. Würschinger & McGillivray (2024) applied it to COVID neologisms, not one persistent ideological token.
3. Operationalizes platform moderation events as causal inflection points in semantic drift	Baele et al. (2023) notes language intensifies following content moderation. Ribeiro et al. (2021) documents user migration but does not measure what happens to a specific word.
4. Tracks redpilled across full ideological spectrum simultaneously	Hagen, Tuters & Wilson (2020) tracked frequency only, stopped at 2020, no semantic modelling. All other studies focus on one community (r/TheRedPill, r/Incels) or map the ecosystem without focusing on one token.

³ Marwick & Caplan (2018) on manosphere language and networked harassment. Lindsay (2022) on the blackpill as ideological framework. McGlashan & Krendel (2023/2024) on community-specific lexical distinctiveness as empirical evidence of interpretive community divergence. [Alexander Kalr Maier](#)

-
5. Corpus extends to
December 2024
-

- | | |
|--------------------------------------|---|
| 6. Proposes theoretical
mechanism | All computational studies (Farrell, Ribeiro, Baele, McGlashan) document what the language does or how toxic it is. None theorises why the same token sustains radically different meanings across communities simultaneously. |
|--------------------------------------|---|
-

Theoretical Framework

The following section establishes the theoretical framework upon which the subsequent data analysis is built. It proceeds in four broad movements. The section opens by situating the study within the established tradition of diachronic semantics, surveying the state of the art across both qualitative and quantitative approaches to semantic change. It introduces [Blank's](#) (1999) typology of semantic change as a foundational classificatory framework and traces the S-curve model as the dominant account of how lexical innovations diffuse through speech communities over time. This historical grounding serves a dual purpose: it establishes the methodological lineage within which the present study operates, and it foregrounds the question of whether models developed for pre-digital language change remain adequate to the conditions of digitally mediated communication.

The section then turns to the technologically specific conditions under which language evolves within human-machine environments. Drawing on the concept of constraint-shaped registers (CSR), it traces historical instances of human-machine linguistic coevolution and derives the defining features of language that has developed in sustained interaction with technical architectures. Three distinct modalities of technologically dependent semantic drift are identified and examined in sequence: digitally modified analogue language, in which pre-existing linguistic forms are adapted to digital constraints; digitally native language as instantiated in computer-mediated discourse (CMD), which emerges from and is irreducible to its digital substrate; and pictorial and semiographic language, encompassing the visual and hybrid sign systems characteristic of online communication. These modalities are contextualized within the broader tradition of semiotics through an examination of digital signs, ambiguity, metairony, and memification as distinctly digital semiotic phenomena.

From this typology the section derives the defining structural and pragmatic features of digitally native lexical items and coevolved neologisms. It examines the dependent use form and the mainstreaming process through which digitally native terms migrate across discursive communities and engages with Critical Discourse Analysis as a framework for interpreting the ideological valences such migration carries. The concept of unlimited semiosis is

introduced to account for the theoretically open-ended proliferation of meaning that characterizes digitally native language, and the semiotics of cyberspace is examined as the broader symbolic environment within which such proliferation occurs.

The section concludes by addressing the methodological implications of this framework for the study of semantic drift in digital corpora. It argues that the unique conditions of digitally native language necessitate an adaptation of historically established methods, and makes the case for a mixed method research design capable of capturing simultaneously the quantitative trajectory of semantic drift, the sociological process of mainstreaming, and the ideological potency that attaches to a term as it traverses discursive communities (i.e diachronic and synchronic analysis). These preliminary contemplations allow the subsequent section to establish *redpilled* and its orthographic and morphological variants as well-suited lexical item for a paradigmatic analysis.

A History of Semantic Drift

In the history of the linguistic tradition the study of semantic change (or shift, or progression) refers to the evolution of a word's meaning over time. The discipline has long engaged with well-documented examples such as “awful” and “nice” ([Geeraerts & Cuyckens, 2010](#)) alongside modern examples that figure more predominantly in contemporary discourse (ex. Queer). So far, historical linguistics has appraised how entire language systems undergo diachronic semantic change; i.e meanings shift across languages, their locales, and regional or social varieties ([Traugott & Dasher, 2001](#)), ([Li & Siew, 2022](#)). Such examples are usually taken to index broader historical processes.

By contrast, the semantic development of individual lexical items whose emergence is unique to, or contingent upon, a specific medium, particularly digital media, remains comparatively understudied. This gap in the research is perhaps unsurprising, given that the dominant methodological frameworks of historical linguistics were devised to account for long-term diachronic processes rather than rapid cycles of lexical innovation ([Radulović, 2023](#)). Yet the proliferation of digitally mediated discourse has produced a range of semantic phenomena that resist straightforward analyses under these established models. Accordingly, this section briefly adumbrates the methods employed in historical linguistics and projects these frameworks, *mutatis mutandis*, onto the new technological environments that structure contemporary digital communication. Rather than repudiating the classical tradition outright, the aim is to interpolate its explanatory tools into a new communicative environment and to nuance semantic methodologies, as has been discussed in prior studies ([Hamilton, Leskovec](#)

& Jurafsky, 2016).

State of the Art

Notwithstanding previous technological limitations, historic semantic analysis aims at portraying meaning not as property internal to lexical items, but as a relational phenomenon shaped by contextual use (Blank & Koch, 1999, here 2013). Previous studies interpolated the convened understanding of diachronic semantic drift, by mapping lexical evolution beyond formal linguistic features, situating terms within broader networks of social practice and communicative pressures (Traugott & Dasher, 2001). It is hence neither fixed nor autonomous but relies on patterned usage across time and context (in this study referred to as iteration). To visualize these multicausal processes, scholarship has traditionally relied on the longitudinal textual evidence, qualitative interpretations, which are combined with distributional patterns observed in historical corpora (Walkden et al., 2023). Changes in collocational behavior, in pragmatic function, and alterations in evaluative or ideological loading are typically taken to index processes of semantic reanalysis.

It starts by defining the analytical features of lexical items (here, tokens) by. To arrive at these preliminary understanding of the object of inquiry, diachronic linguistic analysis necessitates to analyze a term's etymology (1), semantic description (2, qualitative), diachronic semantic mechanisms (3, quantitative), representative and / or anecdotal genre usage (4). Change is identified if systematic reinterpretation keeps occurring, the novel sense is conventionalized, and stably persistent over time. We might take the 1989 study by Simpson et al. as exemplary of this methodology:

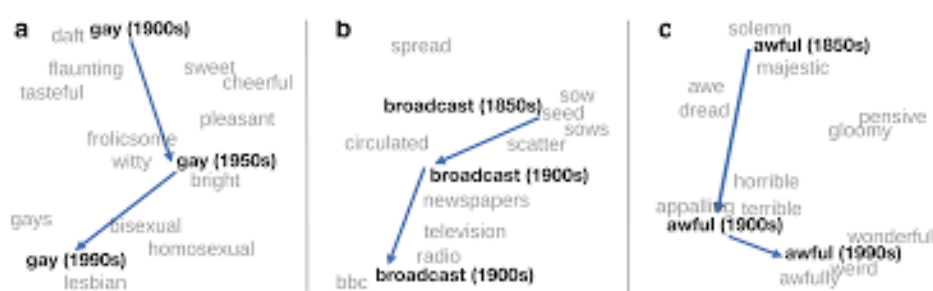


Photo 1 (Hamilton et al., 2016). Citing (Simpson et al., 1989), accessed via <https://cs.stanford.edu/people/jure/pubs/diachronic-acl16.pdf>

Out of these analyses, the research has deduced a plethora of motivations for semantic change (Blank, 1999). Primarily linguistic features such as interpretive openness (fuzziness), socio-linguistic evolutions, including taste (Bourdieu, 1991) or social hierarchies, institutional and non-institutional linguistic prescription, as well as aesthetic and communicative-formal motivations. Some studies have mapped these findings onto digital semantic drift (Solopova

et al, 2025) ([Jaki et al., 2019](#)). However, these studied examples often either attend to the qualitative or quantitative analysis of digital semantic drift, allowing a new body of research to emerge ([Tahmasebi et al., 2021](#), found through FU Berlin). Furthermore, adapted methodologies, specifically apposite to digital semantic drift, will be presented in the following section.

Qualitative Methods

Erstwhile, lexicographic as well as etymological analysis and comparative historical methods were the foundational tools of the discipline ([Campbell, 2013](#)), providing a basis for tracing the formal and semantic evolution of words across time and language families. Corpus-based qualitative analysis extended this reach by enabling in-depth, contextually grounded across large textual datasets ([Hasko, 2012](#)), while Discourse Historical Analysis (DHA) captures the ideological and institutional dimensions of semantic change ([Reisigl & Wodak, here 2017](#)). These approaches are often subtended by thematic frame analysis ([Fillmore & Baker, 2012](#)), thus anchoring meaning within structured background knowledge, and sociolinguistic ethnography or field surveys ([Milroy & Gordon, 2002, here 2006](#)). It underwrites and grounds semantic variation in the lived social contexts of speakers. These approaches remain pertinent in the digital realm, insofar as they sensitize the researcher to ideological and social context; even where digital corpora demand adapted methods.

While this thesis acknowledges amelioration and pejoration of meaning ([Ullmann, 1962](#)) and the strengthening and weakening of meaning ([Bloomfield, 1933](#)) as historically significant descriptive categories, it treats them as peripheral and not objectively classifiable. When trying to apprehend the semantics of fringe and non-mainstreamed communities, whose epistemologies remain often difficult to access and to constitute, it might be worthwhile to omit said approaches from the analysis. Digital semantic drift accepts the concurrent existence of multiple meanings, resisting neat directional categorization. As one corollary of these apprehension this thesis will use [Blank's \(1999\)](#) typology of semantic change⁴:

Metaphor	Change based on similarity between concepts, e.g., mouse "rodent" to "computer device".
Metonymy	Change based on contiguity between concepts, e.g., horn "animal horn" to "musical instrument".

⁴ Blank considered it problematic to include amelioration and pejoration (as in Ullmann, 1962) and strengthening and weakening of meaning (as in Bloomfield, 1933) as primary categories. According to Blank, these are not objectively classifiable phenomena; moreover, all examples listed under these headings can be subsumed under other categories in his typology, rendering them redundant.

Synecdoche	A type of metonymy involving a part to whole relationship, e.g. "hands" from "all hands on deck" to "bodies"
Specialization of meaning	Downward shift in a taxonomy, e.g., corn "grain" to "wheat" (UK), to "maize" (US).
Generalization of meaning	Upward shift in a taxonomy, e.g., <i>hoover</i> "Hoover vacuum cleaner" to "any type of vacuum cleaner".
Cohyponymic transfer	Horizontal shift in a taxonomy, e.g., the confusion of mouse and rat in some dialects.
Antiphrasis	Change based on a contrastive aspect of the concepts, e.g., perfect lady in the sense of "prostitute"
Auto-antonymy	Change of a word's sense and concept to the complementary opposite, e.g., bad in the slang sense of "good"
Auto-converse	Lexical expression of a relationship by the two extremes of the respective relationship, e.g., take in the dialectal use as "give".
Ellipsis	Semantic change based on the contiguity of names, e.g., car "cart" to "automobile", due to the invention of the (motor) car.
Folk-etymology	Semantic change based on the similarity of names, e.g., French contredanse, orig. English country dance.

Blank, Andreas (1999), "Why do new meanings occur? A cognitive typology of the motivations for lexical Semantic change", in Blank, Andreas; Koch, Peter (eds.), *Historical Semantics and Cognition*, Berlin/New York: Mouton de Gruyter, pp. 61–90, for formatting reasons this table was copied from www.wikipedia.com but can be equally found in the original source pdf

Quantitative Methods

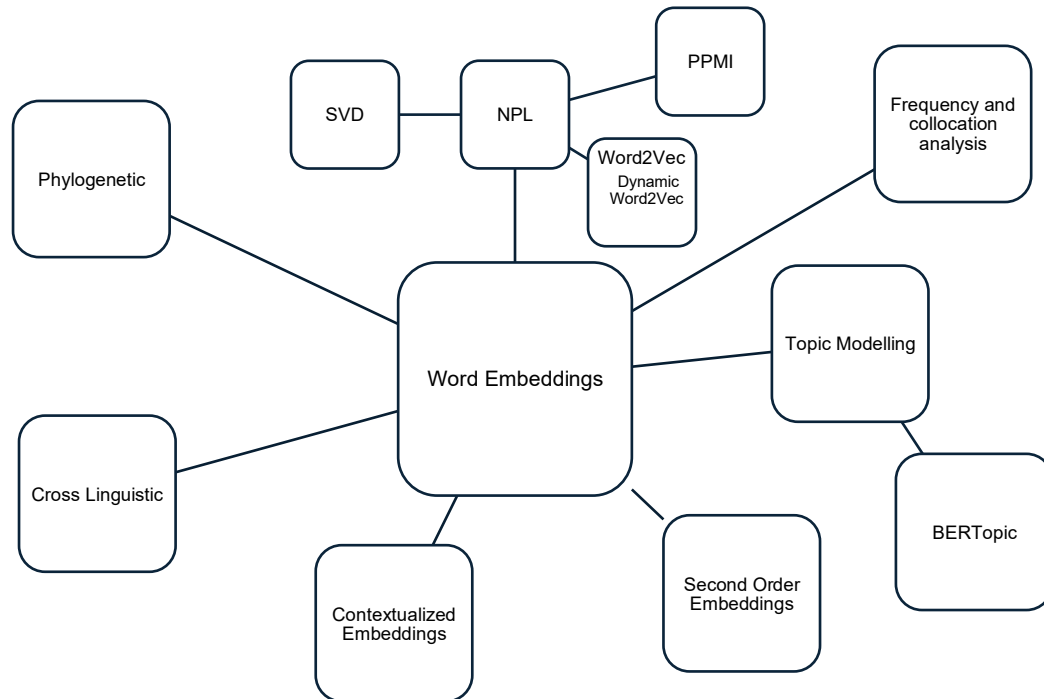


Figure 3 Made with Word Smart Art Design

With the increasing availability of large digital corpora, word embeddings have become the dominant quantitative tool for studying semantic drift, allowing words to be represented as numerical vectors and changes in meaning to be measured as shifts in geometric distance within vector space. The methodological framework developed by [Hamilton, Leskovec, and Jurafsky](#) constitutes exemplary for the most widely adopted approach in the field, within which semantic change is quantified by evaluating several embedding models, including PPMI, SVD, and Word2Vec, against known historical cases of language change across six corpora spanning four languages and two centuries, and applied in the present study to a digitally native corpus, extending its application from long-term historical language change to the accelerated and community-stratified conditions of online discourse ([Hamilton et al.](#), 2016). This adaptation is necessitated by the structural properties of digital language outlined in the preceding theoretical section, wherein meaning shifts occur under unique circumstances.

Several alternative approaches to measuring semantic change have been proposed in the literature and are noted here for comparative context. Dynamic Word2Vec represents a significant methodological development within the embedding paradigm, jointly learning and aligning word embeddings across time periods rather than aligning independently trained

models post hoc, thereby enabling the identification of evolving semantic associations directly within diachronic corpora ([Rudolph & Blei, 2018](#)). Contextualized embeddings, most notably BERT, have further extended the range of available approaches by producing context-sensitive representations through pre-training deep bidirectional transformer models on large corpora of unlabeled text, allowing the meaning of a word to vary across instances rather than being fixed to a single vector ([Devlin et al., 2019](#)). A further development within the transformer paradigm is the sentence embedding approach, in which the entire comment or document is encoded as a single fixed-dimensional vector rather than at the word level, enabling semantic similarity between texts and category descriptors to be measured directly via cosine distance; allowing large corpora to be classified into semantic fields without annotated training data by matching each document to the category description it most closely resembles in embedding space ([Reimers & Gurevych, 2019](#)). It is this sentence embedding approach, implemented via the all-MiniLM-L6-v2 model from the sentence-transformers library, that is employed in the present study for the semantic field classification of the full 780,000-comment corpus. Second-order embeddings have additionally been proposed as an alignment-free alternative, representing a word by its similarity profile relative to a reference vocabulary rather than by its absolute position in vector space ([Eger & Mehler, 2016](#)). This approach is equally reflected in the present study. At a broader level, it has been established that lexical change in digitally mediated communication follows identifiable diffusion patterns shaped primarily by demographic similarity between communities rather than geographic proximity, with social network structure identified as a primary driver of how new linguistic forms propagate across platforms ([Eisenstein et al., 2014](#)).

Frequency and collocation analysis are additionally employed as corpus-linguistic complements, allowing shifts in word co-occurrence patterns to be tracked over time in a manner that is both computationally tractable and interpretively transparent. Topic modeling techniques are further considered for identifying thematic clusters within the Reddit corpus; in a comparative evaluation of LDA, NMF, Top2Vec, and BERTopic on Twitter data, both BERTopic and NMF were found to offer particular utility for short-text datasets, with BERTopic providing strong thematic separation and NMF results more closely aligned with human interpretive judgment ([Egger & Yu, 2022](#)). However, topic modeling alone does not resolve the question of which broader discursive field a given comment belongs to, particularly where semantic fields overlap or where the same term is deployed across ideologically distinct communities. To address this limitation, a sentence embedding classification procedure is additionally employed, in which each comment is encoded as a

fixed-dimensional vector using the all-MiniLM-L6-v2 model from the sentence-transformers library and assigned to the semantic field whose natural-language descriptor embedding exhibits the highest cosine similarity; this approach enables theoretically grounded field classification at scale without requiring annotated training data or supervised fine-tuning ([Reimers & Gurevych, 2019](#)). The present study does not employ approaches such as cross-linguistic or phylogenetic modeling, nor statistical techniques such as representational similarity analysis, as these methods typically require either multilingual datasets or historical time spans not available in the Reddit-based corpus used here.

The S-Curve Model

Previous findings in the studies of language subtended the claim that linguistic innovation initially spreads slowly, accelerates mid-proliferation as it reaches a critical mass of speakers, then slowing towards the end. ([Tagliamonte & D’Arcy, 2009](#)). The now well-established S-Curve model (Altmann et al, 1983, not in bibliography because not accessible through Sciences Po Library; yet cited in [Denison, 2009](#)) ([Labov, 1994](#)) provides a statistical trajectory, “*in which frequencies are cumulative*” ([Tagliamonte & D’Arcy, 2009](#): p. 58) (See Figure 4).

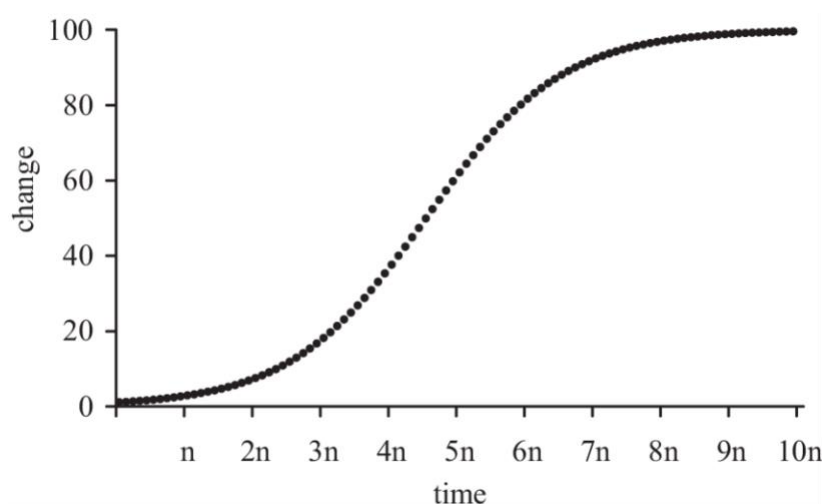


Figure 4 ([Tagliamonte & D’Arcy, 2009](#))

Contra, the often-discussed transmission ([Labov, 2001](#)) and incrementation problems (contemplating each how linguistic change is passed on how it advances), posit the role of imperfect learning ([Halle, 1962](#)) and peer group acquisition within diachronic linguistic change. It ascribes the temporally stacked dynamics of the S-models ([Feltgen et al., 2017](#)) to social group identity ([Labov, 2001](#)) and social network density ([Milroy & Gordon, 2006](#))⁵. In

⁵ I: Speakers recreate the speech patterns of their social surrounding

addition to the randomly distributed neutral drift, research also recognizes the importance of selection bias⁶ and threshold effects⁷ Digital media fosters a transmission ecosystem that equally follows a S-curve propagation model ([Würschinger](#), 2021), and within this light requisites especial recognition of social rather than geographic proximity ([Eisenstein et al.](#), 2014); Our findings also suggest that they may particularly amplify mechanisms of social diffusion, selection biases, and learning thresholds, as algorithmically structured and highly networked online spaces intensify exposure and accelerate the adoption of emerging linguistic forms.

Human-Machine Language Coevolution and Constraint Shaped Registers ***Technologically Dependent Semantic Drift***

Historically, language has always coevolved with its contemporary media. The storied analogy of the *pharmakon* (φάρμακον), re-popularized in midcentury theory by thinkers like Jacques Derrida, encapsulates the inherent tension of writing; that is, the perils of *téchne* (τέχνη) in our modes of memory and argument ([Derrida](#), Barbara Johnson Translation of 1981). Its instable nature has been the subject of the 20th century’s philosophical crisis ([Rorty](#), here 1992); ([Wittgenstein](#), 1921, here 2014); ([Livingston](#), 2017). The history of the literary tradition, from a technological perspective, has therefore been partly contingent on the linguistic capacities engendered by its architecture (i.e its technological coevolution). The standardization efforts of printing machines (Cf. Gutenberg) ([Okrent & O’Neil](#), 2021, pp. 116-146) serve as an emblematic example that has both shaped the tools of lexical expression (codified lexical rules, orthographic correctness) and its proliferation speed ([Nevalainen & van Ostade](#), 2006, pp. 271–311). In its mutation from spoken *lexeme* to expressive written token, the oral tradition of language has increasingly been mediated by these forms of written expression. The Internet may represent the continuation of this process (i.e the centralization of logocentrism, ([Derrida](#), 1967, here 2016)).

The Written Speech Paradox

Therefore, digital semantics requisites a nuanced understanding of how meaning is constructed. Digitally mediated language fosters an environment, in which conversational speech and written modes of expression might occur simultaneously at any given time, yet without the embodied speech elements (such as intonation etc.); this is referred to as the ‘Written Speech Paradox’ or hybrid register ([Tagliamonte & Denis](#), 2008; “[*Instant*

⁶ II :Speakers prefer the easiest to use method within a medium

⁷ III: Speakers may adopt new lexical items after a critical mass has adopted them

Messaging] is firmly rooted in the model of the extant language.”, p.3). Due to the absence of paralinguistic cues ([Kurniasih](#), 2017), speakers have acquired a set of communicative strategies and linguistic expressions, coalescing to what is often referred to in the literature as *Computer-Mediated Discourse* (CMD) ([Herring](#), 2001; [2007](#)) or here *Constraint Shaped Registers* (CSR; see below), and to which diachronic linguistic analysis must attend. Notwithstanding these limitations, it also allows for new linguistic productivity (ex. morphological productivity ([Crystal](#), 2006, “ [P]eople regularly play with the typography or morphology, producing linguistic creations of sometimes virtuoso quality” p.161), particularly with compound words that in spoken English would be hardly grasped), and whose positive role ought to be also recognized by this study. The research constitutes three new linguistic and intrinsically literary forms native to digital communication ([Herring & Androutsopoulos](#), 2015, Chapter 6); analog language patterns used on and thus shaped by digital environments (1), digitally native language (2), and semiotic/ pictographic communication (3).

(1) Digitally Modified Analogue Language

Today, it is a predominant notion that technology accelerates or at least nuances semantic drift ([McCulloch](#), 2019). It prominently features a linguistic economy with its typical abbreviations (see below) as well as register collapse (also bleeding), in which literary forms mix (ex. Internet abbreviations in academic writing). Semantic recontextualization remains the most present linguistic evolution that has reentered oral language ([Thorne & State](#), 2003). One might think of “meme”, “feed”, or “cloud” as prominent examples of how an analog term’s significance may be broadened or metaphorically appropriated within new technological environments ([Hurbrik & Khan](#), 2025).

(2) Digitally Native Language (Computer-Mediated Discourse; CMD)

Additionally, a new language is produced, *aka* digitally native language (DNL). This thesis considers DNL to comprise lexical, syntactic, and discourse forms that emerged organically within digital communication environments. Inevitably, they are shaped by the technical affordances and social dynamics of specific platforms and carry meanings that are hardly intelligible outside of their digital context of origin. A list of historic and adjacent conceptualization remains relevant, such as digital vernacular, Computer-Mediated Communication (CMC) and Internet language ([Herring](#), 2004; 2007), NEAT slang, Netspeak ([Crystal](#), 2001), Chatspeak ([Squires](#), 2010), or the recent Algospeak ([Aleksic](#), 2025).

With the rise of the early Internet, new linguistic forms were apprehensively anticipated, highlighting their profound impact on the legitimately established literary forms within the

liberal script ([Crystal](#), 2001) ([Cystal](#), 2006). Now we might appraise a plethora of linguistic phenomena unique to this form; most commonly neologisms (such as “Selfie”, “Doomscrolling”, “Rizz”, “Redpilled”, “NPC”), which function under collapsed timescale. Others include hashtags as syntactic units ([Zappavigna](#), 2018); or as interpersonal community tools⁸, and the acceleration of viral slang within rapid lexical cycling due to platform algorithms ([Aleksic](#), 2025). These forms often occur as new ways to express the “unsayable”; either due to paralinguistic limitations or content moderation ([Steen](#) et al., 2023).

Despite the general concerns, cyber language has sharply risen and found its way into embodied and every-day speech. These evolutions are natural adaptations to the constraints of synchronous text-based chats without paralinguistic cues ([Crystal](#), 2001, p. 28, and p.37)⁹. Famously, character limits on *X* generate a specific literary style with its own cultural features. This study identifies these *Constraint Shaped Registers* or *Computer Mediated Discourses* as the linguistic premise for digital semantic drift. Prior studies have also shown that specifically these meaning changes can be quantified with statistical models and corpus structures ([Hamilton](#) et. Al, 2016) or that distributional methods may be applied to historical semantic change detection ([Schlechtweg](#) et al., 2024).

Semiotics of the Cyberspace

(3) Pictorial and Semiographic Language

Social media semiotics places special attention to how signs (partly) mediate interactions within CSR or CMD on online platforms. Drawing from the classical study of semiotics as dyadic (Saussure, 1916, here Wade Baskin translation (1959, here 2011)) (Cf. Photo 2), which understands a dichotomous relationship between signifier and

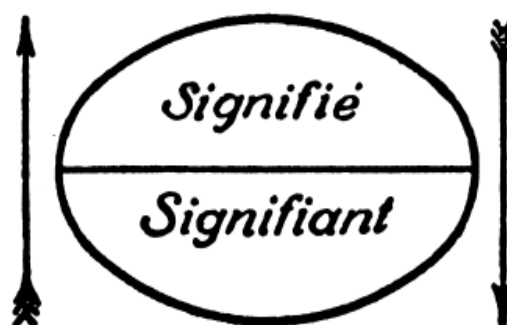


Photo SEQ Photo * ARABIC 2 Taken from Saussure's [Course in General Linguistics](#), 1916

signified, and which later had been nuanced by considering larger cultural artefacts (Barthes, 1957), this field expands (digital) semantic shift by highlighting the potential ambiguity and multi-significance any textual item might contain. Due to its technical features that rely either on written textual containers, expressed through or within code strings, or visual cues (ex. Images, Emojis, Video), i.e a multimodal media system, the Internet has increasingly been

⁸ Zappavigna, M. & Martin, J. (2018) Performing Identity: A Topological Perspective, in *Discourse and Diversionary Justice* (pp. 245–268). Cham: Palgrave Macmillan.

⁹ Accessible via: http://library.megu.edu.ua:8180/jspui/bitstream/123456789/4800/1/2004-Language_and_the_Internet.pdf

recognized as a *semiosphere*; drawing on ([Lotman & Clark](#), 1984, here 2005 pp. 205-229); ([O'Halloran et al.](#), 2013; later too for section on multimodality)¹⁰. Perforce, Internet users have been ascribed both the role of a creators and consumers of said signs, playfully reinterpreting the socio-semiotic resources at their disposal ([van Leeuwen](#), 2004). These so-called conventionalized signs (emerging from interpretive communities) are created to shortcut conveying larger ideas, signal group belonging, or avoid automatic moderation (ex. Plausible deniability). While any such word token might become a sign, certain, especially imaginative word creations unique to the Internet, have proven to be larger vessels for digital sense-making (i.e. digitally native neologisms) These might store entire imaginaries and aesthetic expressions within efficiently short or clandestine wording. Qualitative literature elucidates equally that these signs are dynamic actions and artefacts, which vary over time and may be employed differently by various sociolinguistic communities ([Thurlow & Mroczek](#), 2011) (and others).

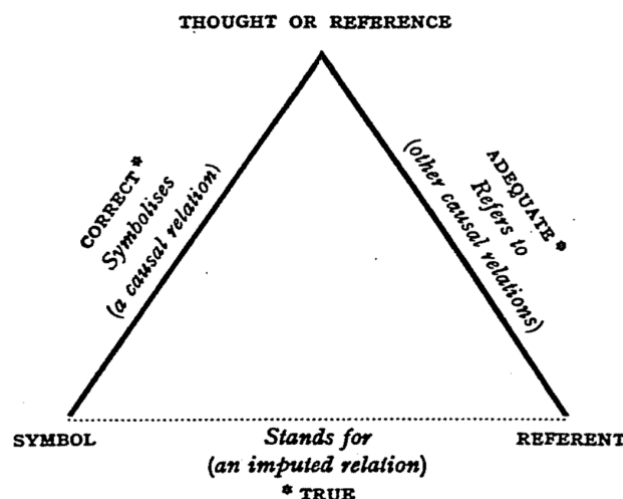


Photo 3 Taken from p. 11, *The Meaning of Meaning* ([Ogden & Richards](#), 1923)

Figure 5

Digital Signs and Ambiguity

While Peirce's triangle (See Photo 2) remains a foundational reference point, poststructuralist readings have further destabilized the assumed stability of the sign. The socio-technical features of the Internet as a *semiosphere* favor multimodality ([O'Halloran et al.](#), 2013), that is the concurrent deployment of text, image, sound, and gesture as co-equal expressive modes, and with it large systems of intertextual reference and an absence of analogue referents. When contemplating digitally native neologisms particularly, this study found it indispensable to attend to the mechanisms by which semantic fields are nuanced through potential

¹⁰ Important to note are also Vladimir Vernadsky's biosphere and Jakob Johann von Uexküll's Umwelt

representations and the intertextual interplay of verbal and non-verbal communication ([Muhammed & Arafah](#), 2022). The resultant signs often require no stable referent in the analogue world, are subject to the timescale problem of accelerated diffusion across communities lacking shared context (Aleksic, 2025), are rendered ambiguous by the reduced availability of paralinguistic and social contextual cues in asynchronous computer-mediated communication ([Walther](#), 1992), and are actively selected for openness by platform architectures that reward interpretive plurality over semantic closure ([Gillespie](#), 2014). Absent tonal anchors, such signs migrate toward the open signifier model ([Chandler](#), 2007, pp. 78-83); ([Laclau](#), 1996). Digitally native neologisms thus require no pre-existing referent in the material world; the sign may accrue meaning entirely within the digital *semiosphere*. This hyperreal condition (Baudrillard, 1981, here 2007), wherein the simulacrum supplants the real and produces signs whose chain of signification terminates not in a stable referent but in further signs, constitutes simultaneously the major structural obstacle and the generative potential of this new form.

Precisely decisive within this condition of referential instability, multimodality functions not as a separate phenomenon but as the operative mechanism through which that instability is produced and sustained. Since any given comment or post may equally be met with a textual or visual reply, the semiotic interplay between these modes is at the root of semantic destabilization. This dynamic gives rise to what may be understood as a verbal-nonverbal continuum, in which the traditional boundary between linguistic and pictographic representation progressively dissolves. Two critical instances of modal transgression emerge from this continuum: the verbalization of the non-verbal, whereby a visual sign (ex.g. meme, emoji, or GIF) is converted into or responds to a verbal utterance; and the visuo-textualization of the verbal, whereby a lexical item acquires pictographic or imagistic properties. Word and image thus become co-dependent visual representations within a broader collective act of sense-making ([Kress & van Leeuwen](#), 2006). Take the following matrix as an example, with ‘Verbal-Nonverbal’ and ‘Nonverbal-Verbal’ representing instances of semantic multimodality.

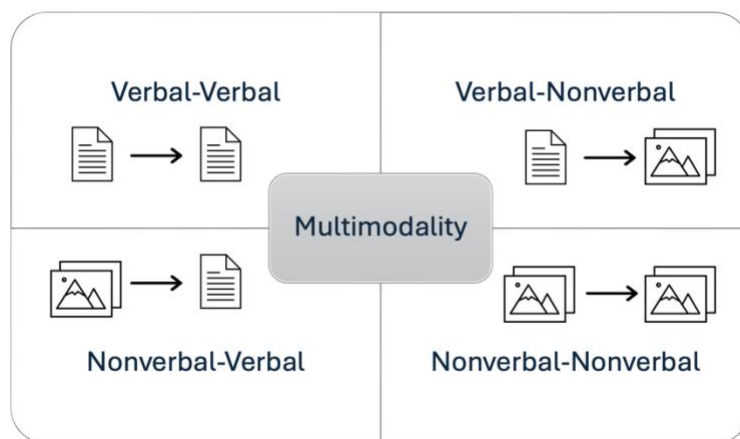
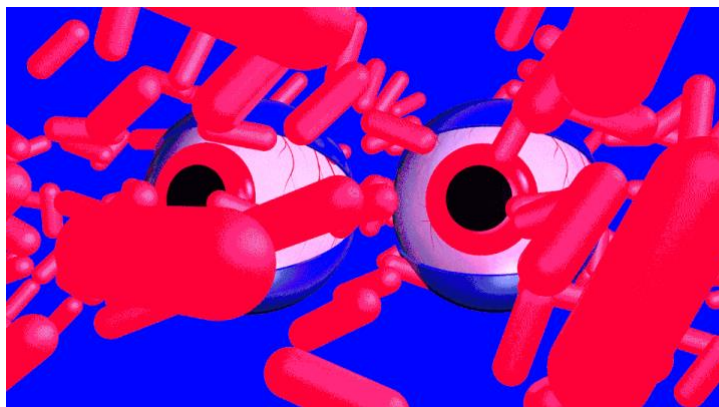


Figure 6 created with Word SmartArt (O'Halloran et al., 2013) & (Hasyim & Arafah, 2022) Having established the sign architecture of the Internet, which relies on layered, multimodal, intertextual, semantics fields without stable analogue referents, said signs may efficiently circulate through various communities due to their capacity to compress large meaning into minimalized digital surfaces. These sociographic communicative “totems” (ex. Memes) have filled an important paralinguistic gap in digital communication and are often politicized (Cannizzaro, 2016); (Shifman, 2014). Their layered meaning, accumulated through repeated multimodal and intertextual use, makes signs politically and cultural “loadable” and a preferred short-cut expressive form. Once again, their creation and consequent diachronic evolution is accelerated and compressed within CMD/CSR (Cf. Timescale Problem (Aleksic, 2025)).

The political productivity of sociographic tokens originates in their capacity to construct a shared imaginary and allow plurality of meaning. Additionally, they often circulate faster than one meaning can stabilize (Würschinger & McGillivray, 2024). Inasmuch as they allow groups to signal belonging and construct larger mythologies, opening signifiers (i.e those that become open signifiers due to high iteration and circulation, hence reappropriation) may also engender coalition building. Thus, interpolating multiple groups at once without requiring a shared stable meaning, over time these might converge or diverge. The possibility of simultaneous meanings destabilizes the sign and its semantic field further, making it hard to pin down one definitive mode of deployment. Therefore, regulatory bodies have struggled to pin down said signs and their potential political messaging without superimposing one fixed interpretation; this thesis will refer to these phenomena as resistance to critique. Furthermore, this openness of meaning allows users to claim plausible deniability. While one user might deploy certain sociographic tokens solely as a linguistic artefact, others might find it to be fit to convey larger ideologies that otherwise would not be spelled out

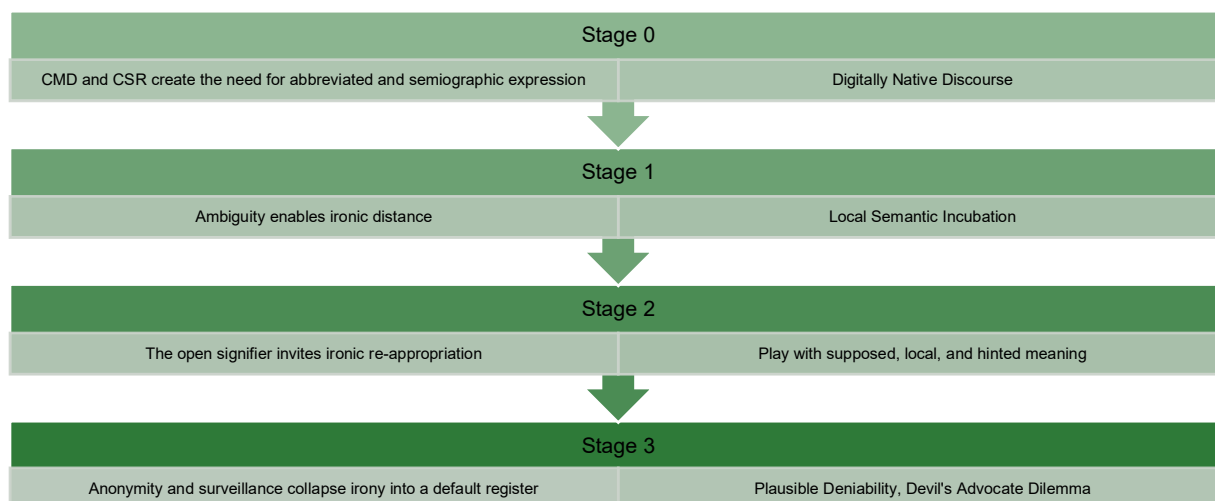
directly. The absence of semantic authorities has been described as a potential pathway to ideological polarization online, while the engagement-optimized platform architectures often reward this interpretively open content and their inherent conflict ([Sunstein](#), 2017). Within this process, memification is the primary cultural technology through which sense chargeability is built and deployed. Due to its innate tension between plausible deniability and political messaging, these forms find larger aesthetic mirrors in its dominant rhetorical mode, *id est* meta irony.

Therefore, to understand digital semantic drift, one must attend to the potential semantic plurality, the playful use of semiotic resources brings to computer mediated discourse. As societal participation becomes increasingly digital, and identity formation mediated through self-expression, consumerism, and entertainment, social media semiotics ascribes large political productivity to these meta-layered signs. This recognition prompted a growing field of research, highlighting the semiotically charged nature of these linguistic phenomena ([Cannizzaro](#), 2016).



*Photo 4 ([Tiffany](#), 2021)
Metairony & Memification*

Ostensibly, fixed image-text constructions that perform communicative acts with a shared cultural meaning, such as memes, may play vital roles within CSRs ([Grundlingh](#), 2018, pp. 147–168) and serve as ironic and/or politically charged totems. They hence relate, inasmuch traditional semiotics does to semantic drift, to the present analysis and must be attributed careful attention, since they reinforce both the meta ironic and political dimensions seen within social media semiotics. Furthermore, they expand the form through visual and in-group signifiers. Drawing from previous analyses, this research understands the ironic nature of the digital *semiopshere* as following:



*Figure 7 Stages of Irony in CSR; made with Word SmartArtDesign
The Dependent Use Form*

Whenever speakers or in this case users interact with signs their relationship to its significance is constructed within their socio-linguistic environments rather than passively received. As Peirce's triadic model classically articulates, signification is not inherent to the sign itself but is generated through the interpretive act of the one who encounters it¹¹. To fully grasp and employ a symbol the speaker must understand navigate three simultaneous relationships; i. e thought-to-symbol (correct), thought-to-referent (adequate), and symbol-to-referent (true). A use of a sign is adequate when it successfully performs its communicative function within a given community; it is true when the symbol accurately corresponds to its referent; and it is correct when it conforms to the normative conventions of the interpretive community that governs the sign (Ogden & Richards, 1923). Any such deployments need only at least two of these relationships to be valid, making the coherence and judgment of meaning governed by normative practices within a community that constantly renegotiates its meaning (Fish, 1980). Irony may be used in an exploitive manner to this dissociation: a speaker may deploy a sign adequately, while simultaneously disavowing its truth claim and pleading ignorance of correctness norms. The gap between adequate and correct is the structural space of plausible deniability.

	Fringe	Mainstream
<i>Adequate</i>	High: sign does its job precisely, even if memeified or ironic	Low: sign misfires or dilutes, might increase if clear political conveyance crystalizes; <i>plausible deniability</i>

¹¹ <https://www.cs.princeton.edu/~chazelle/courses/BIB/semio2.htm>

<i>True</i>	Asserted internally, even if referent is only intertextual, contested externally (cryptic)	Contested from all side; unclear if sign is representing an object, original media, or first digital reappropriation (ex. g. meme)
<i>Correct</i>	Tightly policed internally, <i>per se</i> within community-dependent uses. (“ <i>The joke lands</i> ”)	No authority left to police, <i>plausible deniability</i>

Mainstreaming Process

In the literature, there is now broad consensus that digitally mediated communicative environments facilitate the entry of certain lexical items into mainstream discourse (see S-Curve Model); at times even facilitated by ironic detachment and functioning as a communicative strategy for fringe ideological groups (ex. Dog whistle). Semi-radical spaces often operate as gateway environments for these ideologically charged neologisms ([Scrivens & Gaudette, 2024](#)). Within these spaces, linguistic forms may evolve toward increasingly covert or indirect expressions of hostility, demonstrating both semantic flexibility and a high degree of adaptability across contexts ([Jaki et al., 2019](#)). The gradual normalization of such expressions reflects the broader process of mainstreaming, a well-documented phenomenon in communication studies in which marginal ideas gain broader visibility and acceptance through iterative circulation.

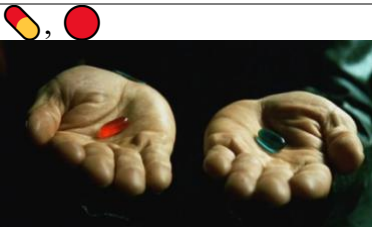
The Semiotics of Mainstreaming

Retaining the proviso that signs on the internet may lack stable analogue referents, and that their meaning remains contextual to the digital communities that employ them, the fringe-to-mainstream divergence outlined above is viewed as a diachronic process; additionally, these semiographic neologisms may accumulate and incubate political meaning over time. In their fringe phase, such signs function as compressed ideological vessels (adequate and correct within the origin community, ironic at the surface as a protective layer against external scrutiny). Ironic humor enhances social and communitarian bonds within these communities, as the correct decoding of polysemic ironic utterances creates a sense of communion, its correctness is policed by tools of active or passive social exclusion. As the sign diffuses, often only in relationship to its semantic broadening through for example becoming sincere vessels to express ideologies and rally support for social and political causes, the ironic register is progressively abandoned in favor of sincere ideological commitment. Yet mainstreaming tends toward a terminal condition: the framing of stereotypes as ironic is used in many cases

to whitewash their ideological content, while over-circulation simultaneously detaches the sign from any single community's authority over its meaning. Once all three criteria are weakened, the open signifier risks becoming an empty one ([Laclau](#), 1996): retaining cultural recognizability while losing the precision that made it politically potent.

Critical Discourse Analysis and Netnographic Approaches

The tension between a term's mainstreaming and its politicization through which many but not exclusively all may become more accessible and sincerely used, is an additional layer previous methods might help to uncover. While the aforementioned corpus linguistics and network analysis bear on the semantic drift of any lexical item, the consequent impact of this migration, especially when contingent upon a political use case should be contextualized. Critical Discourse Analysis (CDA), particularly the three axis approach of socio-cognitive analysis ([Van Dijk](#), 2008) has been essential to understand how language translates into power; [Hertiage's](#) 2023 study on Incel language outlines via netnographic methods, i.e the adaptation of ethnographic methods to digital polities ([Kozinets](#), 2020), how said communities use language to construct gender and race. Furthermore, [Hoebanx](#) (2024; here 2025) analyzes through a mixed methods approach how women employ language in *RedPill* circles such as *r/RedPillWomen* (featured in this corpus). Translating the somewhat rarefied findings of semantic drift analysis, CDA helps to contextualize how a lexical item such as a digitally native neologism might become a semantic container and site for both political organization and violence. The following table provides two examples for a textual and visual signs with the methods of netnographic CDA:

Quantifiable string (Signifier I)	Visual Representation (Signifier II)	Referent (Instable)	Meaning (Signified)	
			General	Adapted
Red-, Blue-, Black-, White-, Pilled (Textual)		The Matrix (1999), pill choice scene	to have been converted to a particular worldview or emotional disposition toward reality	Each pill encodes an entire worldview and emotional disposition toward political reality. Red -> manosphere Black -> Incel Nihilism
Pepe The Frog (Visual)		Cartoonist Matt Furie, Boy's Club webcomic character (2005); a	Internet meme figure: by 2016 sufficiently charged that the ADL added it to its hate symbol	4chan/r/the_donald appropriation as a symbol of the alt-right and Trump movement circa 2015–16; within the manosphere, Pepe

Anthropomorphic

frog + its infinite ideological variants (Sad Pepe, Smug Pepe, REEEEE Pepe, Honkler)

relaxed, apolitical frog saying "feels good man"

database, later partially rescinding this

variants encode specific ideological positions; Sad Pepe = blackpill affect; Smug Pepe = redpill superiority; Honkler = accelerationist nihilism Symbol of Nick Fuentes' Groyper movement, mixing manosphere content with far right neo-Nazis ideologies

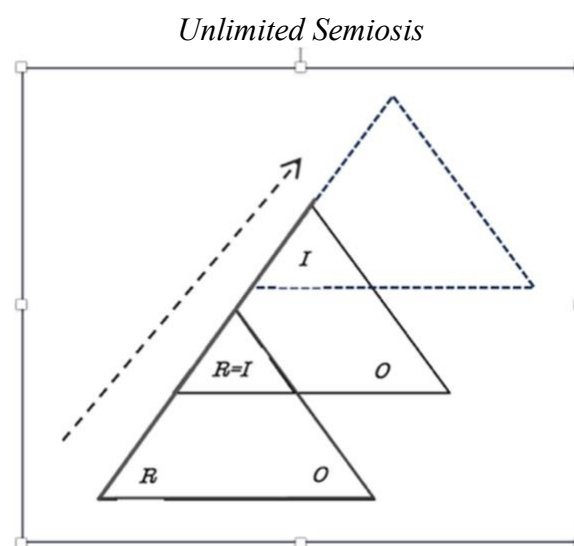


Figure 8 Unlimited semiosis from the triadic model of Peirce (Eco, 1976)

Both memes, 'Red Pill' and 'Pepe the Frog', regardless of their form, adopted a plethora of use cases and with-it associated meanings, illustrating the inherent tension between how such stabilization is both achieved and disrupted in practice. The triadic model contends that the interpretant of any sign becomes a new sign in a possibly infinite process (Eco, 1976). Yet the melability of these words allows the significance to provisionally stabilize through interpretative habits (i.e the habituated attribution of a given signification within a given community). This tension is creating an often-multi-layered discursive realm, consisting of the quantifiable and qualitative (see *State of the Art*) changes a word might experience but also its semiotic representations. In its migratory pattern discourse, communities shape a term's modality, adapting it to their specific discursive needs. A term might henceforth become a new modulation (ex. red to black pill), take on a new meaning through the reappropriation of the former meaning (i.e immediate creation of a new sign; ex. Counter discourse or meta use), and semantic drift (the gradual detachment of any sign from any single

community's authority).

Keeping in mind, the prevalence and trajectory of the S-curve model remain contingent upon the social and communicative environments within which language circulates, the following section outlines the distinctive conditions of digitally mediated language, highlighting the communicative dynamics and platform-specific affordances to which the present study must attend.

Features of Coevolved Language *Neologisms*

Having averred prior understandings of digital linguistic evolution, this thesis turns its eye specifically on the role of neologisms. Digitally native word creations betoken former conceptualizations as mediums of expression, mutual recognition, and larger places of sense-making. These emerge qua the present architectural designs, including accelerated diffusion patterns, group identity in anonymous spaces, the creative inversion of platform constraints, algorithmic selection pressures, and community- driven morphological work. A corollary to this might be the absence of gatekeeping institutions that define meaning, but also these communities' hyper-fragmentation and the memetic reproduction between them, often resulting in memification and post-ironic registers. These perennial considerations, show the intrinsic value of play in digital culture that treats language as a creative medium and on which participation might be contingent upon. Whether these lexical changes achieve critical mass before any perspective force can stabilize or police it meaningfully, remains an open question. Tout court, it has been shaped by, among many other, following linguistic phenomena, which may be applied to this case:

Paralinguistic repair, function	Compensation for the loss of prosodic cues (intonation, rhythm, tone) through non-verbal textual substitutes, i.e punctuation, capitalization patterns, emoticons, and emoji deployed as tone markers.
Graphostylistics, form	The expressive use of typographic and orthographic resources as a semiotic system, carrying affect and emphasis in the absence of spoken modality.
Neologism & morphological	The accelerated coinage and diffusion of novel lexical forms, driven by platform constraints,

productivity	community identity needs, and linguistic play. Includes productive morphological paradigms (e.g. the pill-family suffix) constituting system-level lexical world-building.
Constraint shaped registers (CSR) & Computer mediated discourse (CMD)	Distinctive linguistic forms that emerge as creative adaptations to platform-specific technical constraints (character limits, content moderation, algorithmic filtering), producing coded or compressed expression. The constraint is generative, not merely restrictive.
<i>Memification</i>	The templated variation of pictural and/or textual signifiers through replicable formats. Enables irony, syntactic play, and plausible deniability; fixity of form and instability of interpretation allow meaning to migrate without authorial accountability.
Irony & post-irony	Digital communication structurally favors irony by removing prosodic sincerity cues, producing a default mode of oblique expression. Post-irony describes the resulting state of layered meaning.
Code-switching	The individual-level practice of moving between registers, languages, or linguistic personas within a single interaction or across platforms, exploiting the multiplicity of identities afforded by digital anonymity and community membership.
Register multiplication	The system-level observation that digital environments multiply rather than homogenize registers, enabling the coexistence of multiple linguistic personas and modes of expression unconstrained by a single normative center.
The Written Speech Paradox	Constitutive of the medium, not a deficiency of it.
String Semiotics	The process by which a digital character string acquires semiotic function. Sign-making here is

mediated by platform architecture and community
encoding rather than natural language rules.

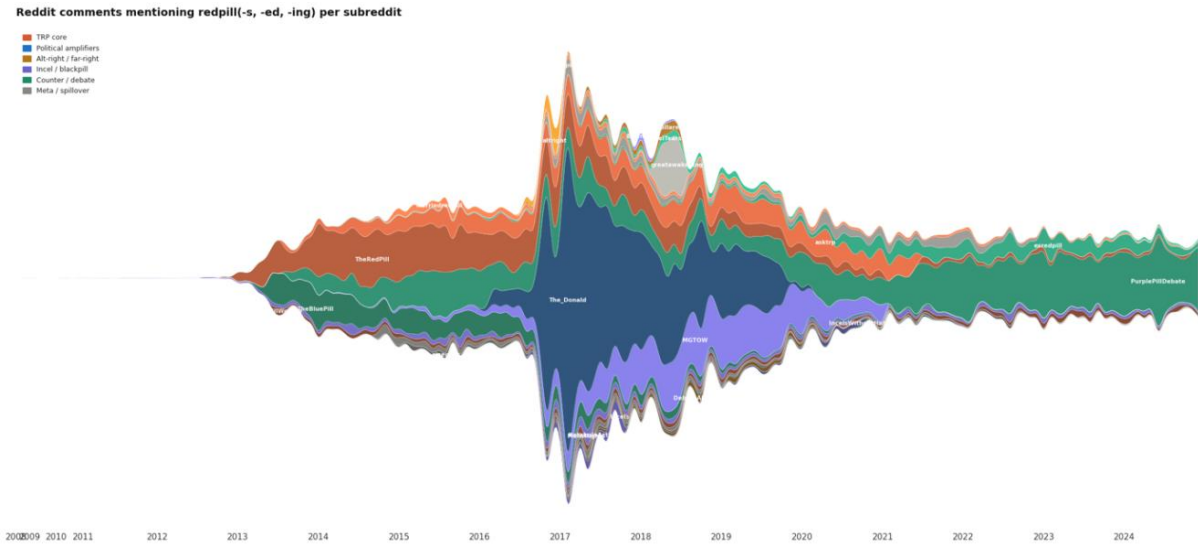
How to Adapt Historic Methods

A diachronic analysis enables the systematic tracing of these meaning shifts by examining how lexical items are deployed across successive temporal stages, within which communities, and within short-circuit drift due to high iteration ([Radulović, 2023](#)). Challenges in diachronic corpus analysis remain the uneven accessibility of historical data and the missing standardization of annotation methods. Equally, neither a uniquely qualitative approach ([Baele et al, 2023](#)), nor a solely quantitative analysis (Cf. Critical Discourse Analysis & Digital Semiotics) can paint all possible modulations a single term might inhabit, nor the imaginaries it might produce. While prior studies of digital linguistic environments have opted for either one or the other, often looking at the specific use within a community (Cf. Existing Studies), the unprecedented study of single lexical item must cash out both methodic paths to also capture a term's post-mainstream afterlife (Particularly, when considering ironic and post-ironic use common in the digital realm). Thus, to showcase both the diachronic change (often expressed in quantifiable data) and its synchronic use cases (established through CDA and qualitative contextualization) ([Kawaguchi et al., 2011](#); recognizing hybrid nature of synchronic variation enabling diachronic change), the uses a mixed methodology.

This approach allows this thesis to treat semantic drift not as a series of discrete changes that renegotiate the meaning of lexical items (ex. g. new dictionary definitions), but as an on-going adaptation process, whereby meanings are incrementally reconfigured through iteration. Within this understanding of semantic drift, semantic stability remains provisional, allowing to the trace meaning variations across temporal intervals and across distinct corpora representing different communicative environments. The following section will briefly outline several methods that may be adapted to digital environments.

Main Part

Case Study Setup



Graph from Dataset 1 Reddit comments mentioning redpill(-s, -ed, -ing) per subreddit

Taking the Redpill

So far, the preceding part has established the theoretical underpinnings of this study, providing the analytical scaffolding necessary for an empirical application. This section provides a justification for the selection of *redpilled* (and its associated variations) as the focal lexical item for this diachronic analysis. In doing so, it addresses several interconnected criteria that together make the term uniquely suited to the aims of this study: etymological tractability and ideological relevance (1), documented cross-community migration (2), semantic adaptability (3), and methodological fit with Reddit as a corpus environment (4). Based on these methodological necessities, a list of criteria is provided to facilitate choosing one specific token.

It will be argued that *redpilled* is not merely a convenient or topical choice, but that its specific linguistic and infrastructural history makes it an analytically productive site for examining how semantic drift and ironic detachment operate within and across digital communities over time. The section further situates the term within the broader context of digital linguistics, demonstrating that its post-cinematic discursive life is partly platform dependent.

Criteria for a Suitable Word Token

In order to establish a clear diachronic corpus of a polysemous and semantically unstable word token, whose textual representation can be normalized across a large, multi-community dataset and traced meaningfully over time (i.e societal and numerical scope), the selected token must satisfy a specific set of linguistic, semiotic, and socio-political criteria

([Androutsopoulos](#), 2011), ([Herring](#), 2007). This study has identified a compressive list of criteria to which any such token should comply. Primarily, the term should be digital native, so emerging from or concurrently to the proliferation of the Internet ([Androutsopoulos](#), 2006). The infrastructural constraints of this space (ex. Platform affordances such as anonymity, iterability, and subcultural concentration) should correlate to the term's emergence. Equally it should be studied within one specific platform environment. To situate it within a broader pattern of political productivity, the term should carry a dense ideological charge from its early beginning and preferably signal affiliation to a specific in-group ([Farrell et al.](#), 2019), ([Ribeiro et al.](#), 2021) ([Farrell et al.](#), 2020). Additionally, prior research should suggest a plausible migratory behavior or pathway (cf. "mainstreaming). It also must be morphologically generative, hence producing a family of derived or compounded forms that allow speakers to articulate ideological variation both within and across communities. The token must be capable of functioning as a sincere ideological statement and as ironically detached, with the possibility of appearing as a mocking gesture or within a camp sensibility ([Sontag](#), 1964, here 2018). These ambiguities and aesthetic configuration, already present in traditional publication fora, allow for the discursive strategy of plausible deniability. These terms often already are present within larger extremism lexicons, proving a validated ideological salience by external researchers ([Baele & Brace & Ging](#), 2024).

Most of these criteria emerge from the structural features of digitally native language, in which lexemes are continually negotiated through ongoing discursive interaction. Within the digital public sphere, meaning-making practices and circulating signs are organized around a set of doxa that differ substantially from those governing embodied discourse, thereby producing a distinct lexicon of legitimate language and taste ([Ignatow](#), 2017). Therefore, the token constitutes a semiotically bound signifier, which is recognizable as itself (as a symbol but also codable string) across all communities it might be employed within, even as the signified shifts ([Saussure](#), 1916, here 2011). It cannot be substituted by ordinary language. Owing to mechanisms of content moderation and filtering, digital communication has produced an increasing number of such semiotically distinct terms, whose meanings are shared and stabilized within specific in-groups ([Prem et al.](#), 2023, pp 481–493). Finally, as the pre-condition of cross-community migration, the token must be formally innocuous enough to circulate in mainstream spaces without triggering immediate moderation, while retaining its ideological charge for the in-group, even after Reddit's 2015 adjustments to community policies ([Åkerlund](#), 2021). *The specific token this study has found to be satisfactory was selected in coordination with Sciences Po's médialab as well as Baele and*

Brace. Redpilled and its variations were chosen especially due to its presence in commonly accepted extremist lexicons.

Why Redpilled and Its Variations?

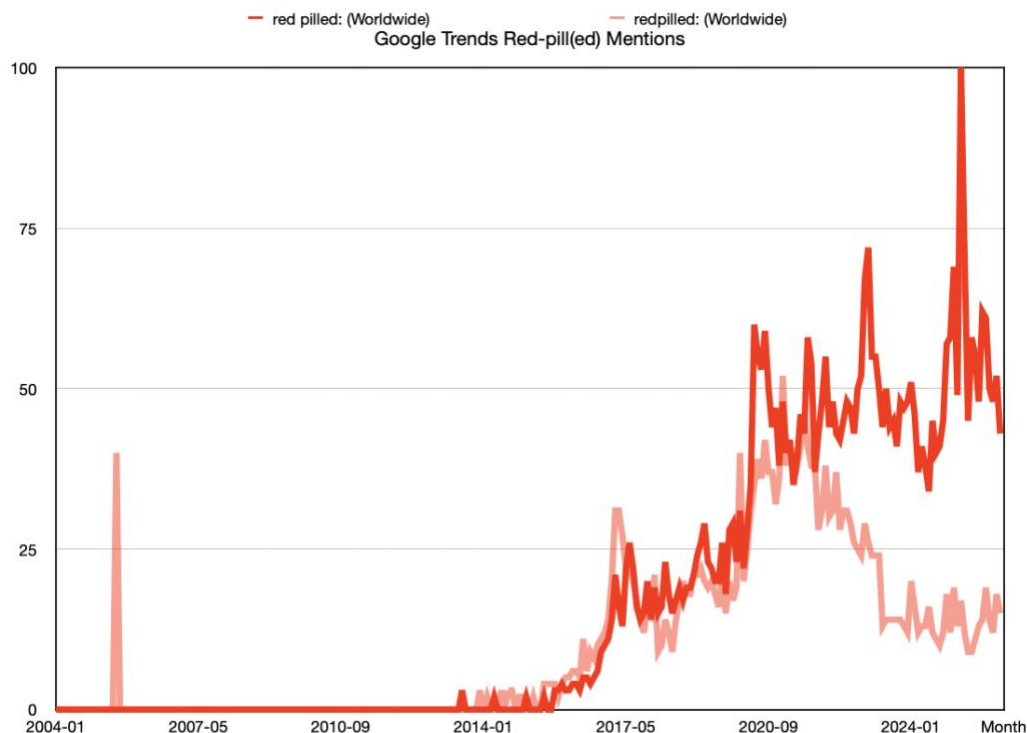


Figure 9 Google Trends of Redpilled, made with Google

Various digitally native subcultures have proven to be incubatory spaces for linguistic innovation. Especially the Incel community and with it the self-proclaimed *manosphere* have stood proxy for these linguistically militating forces. Ipso facto, they have given rise to a large body of creative and digitally unique neologisms that correspond and convey their worldview just like they signal group belonging and lend ways of expression where previously there had been none. Tenable examples are still used today, such as ‘*redpill(ed)*’, ‘*manosphere*’, ‘*lookmaxxing*’, ‘*NPC*’ or ‘*chad*’, which are aimed at obviating thus compressing more lengthy routes of expression. These communities have also favored certain niche lexical items of analogue language that now are disproportionately featured in Internet slang, including *Sigma*, *Alpha*, or *hypergamy*.

Yet the most apposite to the inception of this productive linguistic tradition remains the word *redpill(ed)*. While the Incel movement did not conduce the first appearance of the word, as it is derivative of an external medium (see etymology), its traceability, first adoption, and morphological profile are contingent upon said community’s lifelines. It has also produced a rich morphological family (i.e the Pill Matrix) that allows users playful adaptation of its meaning to local contexts. Today it sits within a productive lexical family: *redpilled*,

bluepilled, *blackpilled*, and morphological variants *-pilling*, *-ed*, *-s*, *-er*. Allowing the dataset to also trace a whole ecosystem of related terms. Furthermore, its traceability is technical as it stands out immediately as a neologism and sign. Unlike words such as ‘*based*’ (=bold, commendable or agreeable person from AVEE, however, can also be just use as the verb to base= ‘This story is based on’) its string features are unique and dependent. While it may take on an equivocal semantic profile its lexical and morphological structure are stable and single use (more than 900k appearances).

Apart from these quantifiable characteristics, the term’s already observed mainstreaming and semantic reappropriation allowed to derive an urgency and special interest. Known as a Matrix-inspired epistemological metaphor, it equally took on the role as collective totem of the *manosphere* and sexual market ideology, an observed politically sincere use during the first Trump administration (*r/The_Donald* era) and conspiratorial chats, as well as a counter-discursive ironic reappropriation and now as a reflexive meta-aware term to describe a political movement ([Tiffany](#), 2021). This digitally native neologism is therefore at the intersection of various ideological movements that all use the term differently, and varying levels of distance from the original meaning. The later presented pipeline normalizes all surface forms (*red pill*, *red-pill*, *redpill*, *redpilled*, *redpillers*, *redpilling*, etc.) into a single token REDPILL using a regex pattern before training Word2Vec.

As of today, far right and other extremist communities deploy so-called dog whistles as coded expressions. These often have dual meanings for different intended audiences, a more acceptable one to the public and a discrete message to the ingroup ([Safire](#), 1988, here 2009); (Lopez, 2015). While a political orientation of *redpilled* is often premeditatively imputed to the Incel-movement, this study complicates the term’s association as a dog whistle. It therefore does not represent a traditional dog whistle but rather a container for ingroup signaling and a totem for ideological sense-making and deliberation.

The term’s discursive life is constituted by a major tension. Throughout its circulation, *redpilled* holds simultaneously sincere, ironic, and post-ironic registers; and even within its sincere use, meaningful distinctions remain between its deployment as a purely anti-feminist manosphere marker, and its use as a vehicle for marrying men’s rights grievance to a broader far-right policy agenda. Over time, the term’s ideological profile has expanded from anti-feminism to anti-elitism to race-based and quasi-Darwinist discourse, accumulating referential fields without shedding prior ones. This makes *redpilled* less a dog whistle in the classical sense than a plurosemantic political vessel, a sign capacious enough to carry competing ideological investments across communities, while retaining sufficient instability

to be permanently contested.

Why Reddit?

The choice of Reddit as the empirical terrain for this study is motivated by a convergence of methodological and structural considerations. It ascribes unique qualities to Reddit in the following categories: archival & technical features (1), platform architecture and moderation history as an analytical resource (2), ideological ecosystem (3), and prior linguistic research (4; Cf. Existing Studies).

(1) At the most practical level, Reddit's archival infrastructure remains accessible and community-oriented (still available to independent researchers post-2023 API restrictions), much in contrast to other social media platforms. The Pushshift.io archive retains extensive historical data stretching back to Reddit's genesis. In line with this study's ambition to construct a diachronic linguistic analysis, inevitably tracking semantic shift across a twelve-year window through yearly Word2Vec models with Procrustes alignments, the methodology requires a large longitudinal and community-partitioned corpus. Furthermore, the Reddit's archival data ecosystem provides diachronic consistencies, and comparatively high posting volume within immediately identifiable, yet niche communities (ex. *r/MGTOW*¹²).

(2) Generally, Reddit is ascribed a specific role regarding platform affordances. In contrast to many of its contemporaries, it allows subreddits, so smaller interest group, to emerge as discrete, yet named, and bounded communities. Keeping in mind, that active participation in any Subreddit implies a premeditative and consented affiliation to the ideological in-group through association. This structural demarcation allows for ideological sampling (ex. *r/Qanon* and *r/thegreatawakening* will mostly share affirmative content on conspiracy theories). Members might join any such group at any time prompting migratory behaviors of said members, but also linguistic features, cultural traditions, and tastes (Ex. overlap between *r/the_Donald* and *r/thegreatawakening*). Reddit's constraint-shaped registers (karma, upvoting, content moderation) equally mirror the sociotechnological environments theorized in the aforementioned part, allowing a unique form of community-specific lexical evolution. Content moderation decision, especially the closing or quarantine of a certain space might prompt large migrations to other, in fact changing their ideological make-up (Copland,2020; indicating that content moderation decisions might not erase radicalized content but push migration into new spaces and new and more clandestine linguistic expressions). In that sense Reddit comes closest to approximating the public sphere

¹² "Men Going Their Own Way": antifeminist, Incel, and manosphere subreddit

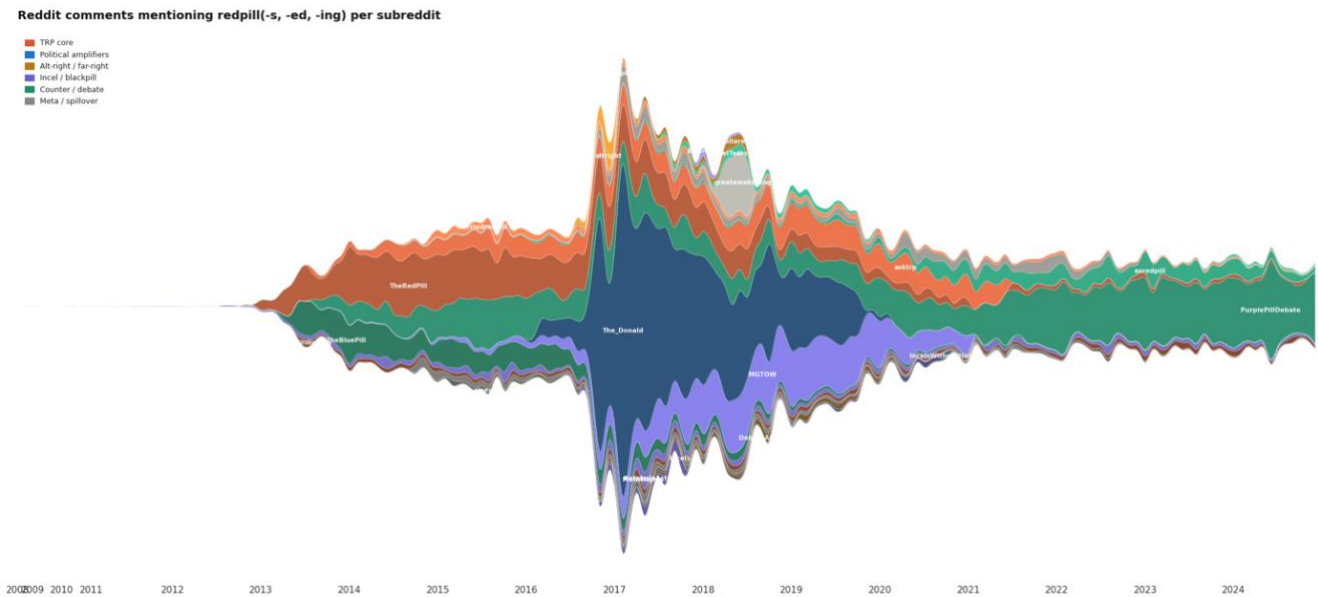
of the analogue life, because it allows the overlap of concurrent discursive spaces. Therefore, a diachronic axis and synchronic axis can be established via time and community within the data set.

Prior research also sustains that the size of subreddits determines its linguistic distinctiveness with medium sized communities being hubs and community size inversely correlating with said distinctiveness ([Lucy & Bamman, 2021](#)): it has equally been shown that small-to-medium communities are important sites of lexical innovation, often dependent on the social relations within these tightly knit environments ([Del Tredici & Fernández, 2018](#)).

(3) The ideological ecosystems under study, such as the manosphere, incels, MGTWO and men's rights activism, blackpill, and conspiracies, plus their respective counter communities (ex. IncelTears, IncelExist) all coexist on the platform, allowing a myriad of semantic coevolution. With its well documented history as an early adopter of digitally native vocabulary, Reddit is often seen as linguistic innovative space, where concurrent to being politically charged, these ideological differences (inevitably due to the CSR) are expressed through the semantic capitalization of a word. While the term *redpill* did not originate on Reddit, it underwent its first large-scale adoption in the subreddit *r/TheRedPill*, promptly undergoing its core semantic development in this and surrounding communities (i.e linguistic innovation as social group formation and political difference). Communities like *r/exredpill* and the feminist *r/ThePurplePillDebate* define their understanding of the word as symptomatic of patriarchal structures and therefore in stark opposition to the *r/TheRedPill*. Extensive prior research on Reddit provides theoretical grounding and comparability (Cf. Part Existing studies).

(4) Within all these reasons, Reddit provides a great breeding ground for linguistic innovation, awarded special attention by contemporary research. With this study's specific focus on digitally native neologisms, Reddit allows to recognize its in communities aka. subreddits to be viewed as innovations spaces or communities of practice with distinct linguistic identities. In the light of the outlined hypothesis, this study allows to understand how the same digitally native neologism carries different semantic associations across communities ([Lucy & Bamman, 2021](#)), and how it is shaped by the structures of both its diffusion patterns and the structures of the subreddits that incubate it ([Würschinger & McGillivray, 2024](#)). A mixed method approach will provide and unprecedented contextualization to this drift.

Data Collection Corpus Building



Having argued, *inter alia*, for the deployment of *redpill* as a pertinent case study through which to trace a word's semantic trajectory within constraint-shaped registers, a subsequent dataset was constructed to lend empirical substance to some of the still inchoate assumptions surrounding redpilled's semantic life. The pipeline broadly follows [Willaert et al.'s](#) methodology in "*A Tool for Tracking the Propagation of Words on Reddit*" (2021), adapting its data-driven architecture to the diachronic and cross-community ambitions of the present study. Qua Reddit's Pushshift.io, which remains accessible to academic researchers, all subreddits registering a minimum of ten monthly mentions of the target term were extracted ($N \geq 10$). Due to the circumscribed availability of Pushshift.io, all subsequent data was retrieved via Academic Torrent as .zst files, amounting to approximately 100 GB of compressed archival material¹³. Figure 7 displays the term's frequency distribution across subreddits and evinces three discernible phases of its migratory pattern. These include 2012-2015 (inception), 2016-2020 (consolidation/mainstreaming), 2021-2024 (post moderation). These can be broadly aligned with the US-American legislative cycles and content moderation decisions on Reddit (see below). The 2020 [OILab](#) study, "*Reactionary Wokeness: How Redpilling Became a Thing on Reddit*" by Sal Hagen, Marc Tuters, and Jack Wilson, returned cognate distributional data for the period within its temporal scope, though without the application of Word2Vec analysis to which the following parts will attend. The analysis returned the following subreddits as principal hubs of redpill usage:

¹³ Contact for access

#	subreddit	comments	group	label	status
1	r/PurplePillDebate	218,139	contested debate / counter-discourse	Neither pole; structured argument space	active
2	r/The_Donald	173,957	political radicalization	Trump-era political awakening	banned Jun 2020
3	r/TheRedPill	149,734	manosphere origin	Foundational TRP community	quarantined Sep 2018
4	r/asktrp	79,753	manosphere origin	TRP newcomer Q&A	quarantined
5	r/MGTOW	79,276	manosphere origin	Male separatism	banned Aug 2021
6	r/TheBluePill	58,612	counter-discourse	Parody / ironic counter	active
7	r/exredpill	50,939	counter-discourse	Deradicalisation / exit	active
8	r/RedPillWomen	30,344	manosphere origin	Gender-inverted TRP adoption	active
9	r/MensRights	23,727	contested debate	MRA overlap; pre-TRP origin	active
10	r/conspiracy	21,740	mainstream / meta	Conspiracist awakening frame	active
11	r/marriedredpill	17,077	manosphere origin	Domestic TRP application	active
12	r/KotakuInAction	9,877	contested debate	GamerGate bridge community	active
13	r/greatawakening	9,345	political radicalization	QAnon & 'Great Awakening'	banned Sep 2018
14	r/niceguys	7,528	counter-discourse	Ironic mockery of TRP logic	active
15	r/IncelTears	6,394	counter-discourse	Anti-incel counter-discourse	renamed

16	r/4chan4trump	6,080	political radicalization	/pol/ to Trump pipeline	banned
17	r/milliondollarextreme	4,494	political radicalization	Alt-right irony / comedy	banned Sep 2018
18	r/DebateAltRight	4,491	political radicalization	Explicit alt-right debate	banned
19	r/Incels	4,072	Incel / blackpill	Incel radicalisation space	banned Nov 2017
20	r/altright	3,235	political radicalization	Explicit white nationalism	banned Feb 2017
21	r/IncelsWithoutHate	2,507	Incel / blackpill	Moderate incel successor	banned
22	r/againstmensrights	1,954	counter-discourse	Anti-MRA counter-discourse	active
23	r/kotakuinaction2	1,084	Contested debate	KiA splinter community	active
24	r/Braincels	940	Incel / blackpill	Blackpill / incel crossover	banned Sep 2019
25	r/IncelExit	896	Incel / blackpill	Incel deradicalisation	active

To account for the volume thresholds per year required to construct a meaningful vector space, as well as for constrained technical capacities, the data set relies on a multistep extraction process. Through a Python-based extraction pipeline, a keyword scanner was deployed on the decompressed files¹⁴, normalizing a regex covering all redpill variants, including closed, hyphenated, spaced, inflected, and recognizably misspelled forms. All variants were normalized to the token REDPILL for subsequent Word2Vec training. This method returned a corpus of almost one million comments. The dataset stores each entry with subreddit, redditor, timestamp, comment text, matched surface form, and an index number. The 966,195 comments were subsequently audited (cf. Appendix: Data Audit), with bots, spam, and content moderation messages removed through username pattern matching and supplemented by manual inspection of high-frequency posters. Deleted

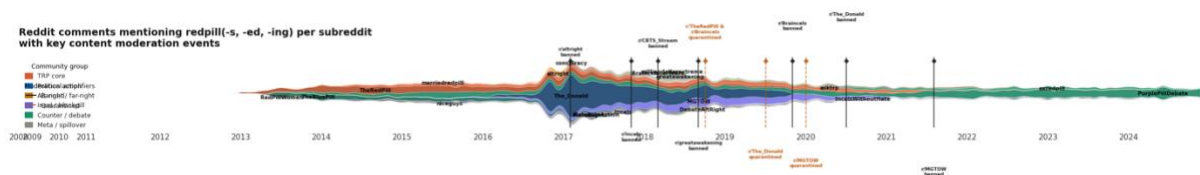
¹⁴ REDPILL_PATTERN = re.compile(r'\bred?s*-?\s*pill(?:ed|ing|er|ers|s)?\b'; re.IGNORECASE);

comments were preserved with author field marked [deleted]. The following figure displays the final mention frequency over time of the audited data set. Inception follows the creation of *r/TheRedPill* and accelerates markedly during the 2016 US presidential election cycle, peaking in mid-2017 before entering a gradual decline.



Graph from Dataset 2 Monthly Comment Volume with "redpilled" mentions

The term's declining deployment is largely explicable by the content moderation decisions taken by Reddit after 2017. Major incubation hubs were subsequently shuttered and their userbases dispersed off platform (Ribeiro et al., 2021). After 2022, deployment of the term remains mostly confined to counter-discourse and or manosphere-adjacent spaces. A structural caveat attends this periodization, however: the approximately 182,000 removed entries cluster disproportionately within high-moderation communities, namely *r/TheRedPill*, *r/Incels*, and *r/Braincels*, precisely during their peak and pre-ban periods. The corpus therefore likely understates both the volume and the discursive intensity emanating from those communities during the term's most ideologically charged phase. This may be intrinsic to the Pushshift.io infrastructure and is not correctable post hoc; accordingly, it is acknowledged here as a constitutive limitation of the dataset rather than a methodological failing of the present study. It furthermore raises ethical questions about the implication of such moderation techniques to the storage and quality of digital archives.



Graph from Dataset 3Reddit comments mentioning redpill per subreddit with key content moderation events, also see table

Methodology and Preliminary Results

Semantic Drift Analysis

To trace the semantic trajectory of redpill across the study period, a discrete Word2Vec model was trained for each calendar year from 2012 to 2024 employing a skip-gram architecture with a context window of five tokens. Prior to model training, all orthographic and morphological surface variants of the target term were normalized to the single placeholder

token REDPILL, and all pill-family tokens were systematically excluded from the surrounding context window to preclude the embedding space from anchoring on intra-community lexical conventions rather than on the broader distributional context of the term. To render the thirteen independently trained embedding spaces mutually comparable, each annual model was aligned to a shared reference space via orthogonal Procrustes rotation utilizing the 2012 to 2014 *r/TheRedPill* corpus as the alignment anchor, the period corresponding to the term's documented point of origin within the manosphere. It bears emphasis that this alignment procedure establishes a common coordinate system across all annual models but does not in itself fix the baseline as the comparator for all subsequent measures. Semantic drift was subsequently operationalized through three complementary measures calibrated to capture distinct dimensions of lexical change. Cumulative drift ($D_{base \rightarrow t}$) directly leverages this baseline by quantifying the cosine distance between the 2012 to 2014 reference vector and the Procrustes-aligned vector for each year t , yielding a longitudinal trajectory that indexes the degree to which the word's meaning has departed from its point of origin at any given moment in the study period.

(1) Cumulative drift: REDPILL vs. 2012–14 baseline

$$D_{base \rightarrow t}(t) = 1 - \cos(v_w^{(ref)}, v_w^{(t)})^{15}$$

$v_w^{(ref)}$ = Procrustes-aligned vector of REDPILL in the 2012–14 *r/TheRedPill* baseline

$v_w^{(t)}$ = Procrustes-aligned vector of REDPILL at year



Graph from Dataset 4 Cumulative Semantic Drift of "Redpill," 2012-2024

Cumulative drift score of the term *redpill* relative to its 2012-2014 baseline vector, computed as the cosine distance between the Procrustes-aligned annual embedding and the reference embedding derived from the founding *r/TheRedPill* corpus. A score of 0 indicates identity with the baseline; 1 indicates maximal divergence. The observed trajectory (red line) shows rapid departure from the baseline through 2013-2016, coinciding with the term's uptake beyond the manosphere into political and conspiratorial discourse contexts. From 2016 onward, drift

¹⁵ All equations were made with Word Equation

scores plateau and stabilize in the 0.44-0.52 range, suggesting that while the term underwent substantial semantic displacement during its early diffusion phase, its distributional meaning reached a new equilibrium rather than continuing to diverge indefinitely. The dashed purple line projects the drift trajectory forward from 2015 at the 2012-2014 growth rate, and the shaded gap between observed and projected values indexes the degree of semantic stabilization relative to that counterfactual.

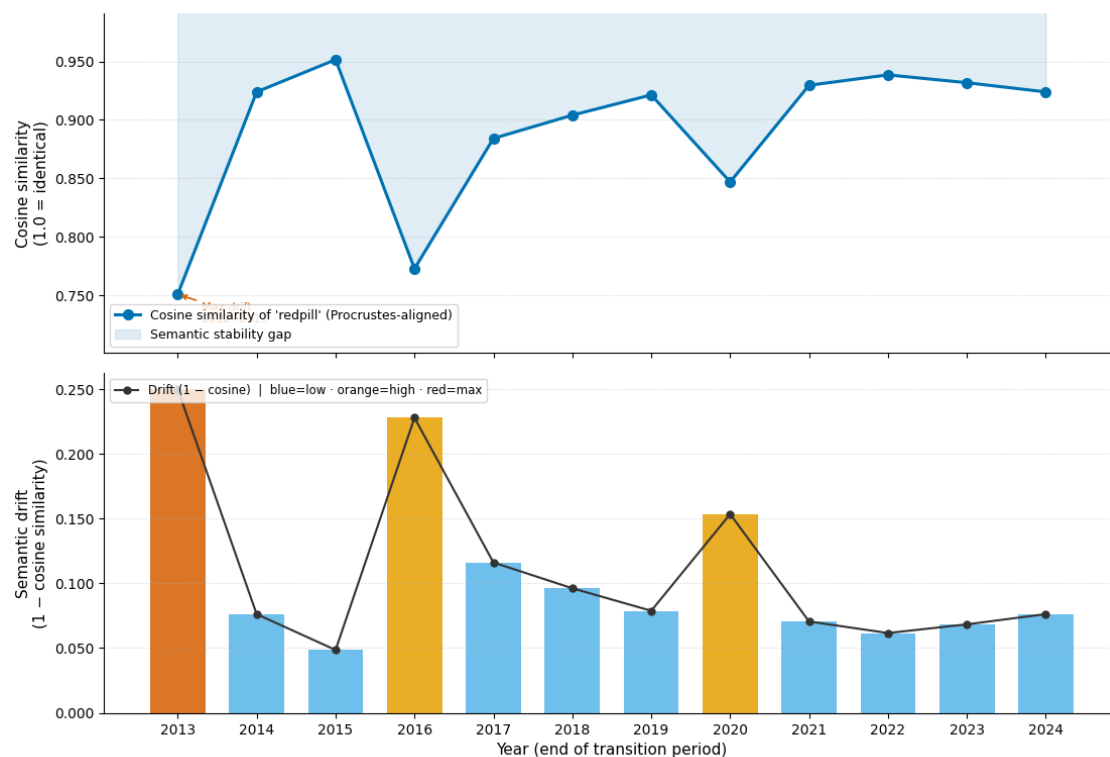
Incremental drift ($D_{t-1 \rightarrow t}$) instead quantifies the cosine distance between the aligned vectors of consecutive years $t-1$ and t , isolating the magnitude of semantic change occurring within each discrete annual interval and thereby locating the precise temporal coordinates of lexical rupture independent of cumulative displacement from a fixed reference point.

(2) Incremental drift: year-on-year change

$$D_{t-1 \rightarrow t}(t) = 1 - \cos(v_w^{(t-1)}, v_w^{(t)})$$

$v_w^{(t-1)}$ = Procrustes-aligned vector of REDPILL at year $t - 1$

$v_w^{(t)}$ = Procrustes-aligned vector of REDPILL at year t



Graph from Dataset 5 Incremental Semantic Drift of "Redpill," Year-on-Year, 2013-2024 Year-on-year incremental drift of the term redpill, operationalized as the cosine distance between Procrustes-aligned embedding vectors of consecutive annual models ($D_{t-1 \rightarrow t}$). The upper panel plots the inverse measure - cosine similarity to the preceding year's vector - while the lower panel isolates drift magnitude with bars color-coded by intensity (blue=low, orange=high, red=maximum). Three pronounced peaks are visible: 2013, reflecting the term's earliest distributional instability as it began spreading beyond its core community of origin; 2016, coinciding with its rapid politicization in the context of the Trump campaign and the broader alt-right media ecosystem; and 2020, suggesting a further moment of lexical

disruption likely associated with the convergence of pandemic discourse, QAnon mainstreaming, and renewed political polarization during the election cycle. The post-2020 period shows markedly reduced year-on-year drift, with values consistently below 0.10, consistent with the stabilization pattern observed in the cumulative drift measure and indicative of a term that has reached a semantically diversified but relatively settled distributional state. Finally, the Global Drift statistic synthesizes all annual displacement scores relative to the baseline into a single weighted summary measure across the full 2012 to 2024 period, scaling each year's contribution proportionally to its relative comment volume within the corpus and thereby ensuring that periods of heightened discursive activity carry commensurate analytical weight (Formula 3). Taken together, these three measures enable a principled analytical distinction between when semantic displacement occurred, at what rate it intensified within any given annual interval, and how substantial the aggregate departure from the word's point of origin proves to be across the study period.

$$\text{Global Drift}(t) = \frac{1}{|T|} \sum_{t \in T} W_t \cdot 1 - \cos(v_w^{(ref)}, v_w^{(t)})$$

T = set of all years in the corpus: {2012, 2013, ... , 2024}

W_t = relative comment volume in year t ; years with more redpill usage contribute more to the overall score; calculated as $W_t = n_t / N$ where n_t = number of comments in year t and N = total comments in corpus

$v_w^{(i,t)}$ = Procrustes-aligned Word2Vec vector of "redpill" in subreddit i at year t

$v_w^{(j,t)}$ = Procrustes-aligned Word2Vec vector of "redpill" in subreddit j at year t

Nearest-Neighbor Extraction

In addition to tracking the semantic trajectory of the target token itself, the study sought to characterize the broader conceptual vocabulary with which *redpill* became associated across the study period. To this end, the 100 nearest-neighbor tokens to REDPILL in the Procrustes-aligned embedding space were extracted for each calendar year using cosine similarity (Cf. Appendix), yielding an annual inventory of the words most distributionally proximate to the target term in that year's discursive context. These raw nearest-neighbor lists were subsequently subjected to a systematic manual audit designed to isolate conceptually meaningful vocabulary from lexical noise. Tokens were excluded from further analysis where they constituted proper nouns, individual usernames, subreddit-specific platform jargon, or *hapax legomena* unlikely to reflect stable semantic associations rather than incidental co-occurrence patterns. Only tokens referring to generalizable concepts, whether ideological, epistemic, affective, or cultural in nature, were retained for subsequent analysis. This filtering procedure ensured that the resulting concept word inventories reflect meaningful shifts in the associative field of *redpill* across time rather than artifacts of platform convention or community-specific naming practices. The audited concept word lists for each year thus constitute the primary empirical basis for a dictionary approach to capture the neighborhood drift and semantic field classification.

***Dictionary*¹⁶**

Matrix / Truth-seeking frame

truth, wake, woke, awaken, awakening, reality, illusion, matrix, simulation, see, seeing, know, knowing, hidden, aware, awareness, enlighten, enlightened, enlightenment, discover, discovered, discovery, lies, lied, blind, blindness, sheep, sheeple, ignorant, ignorance, knowledge, consciousness, veil, perception, deceived, deception, propaganda, brainwashed, asleep, open eyes, unplugged, unplug, free yourself, question everything, uncomfortable

¹⁶ Word doubling to include Standard American English and British, Canadian, Australian *et cetera* spelling

truth, hard truth, real world, seeing clearly, cognitive dissonance, normie, normies, bluepill, blue pill, bluepilled, epistemic, wake up, eyes open, hidden truth, forbidden knowledge, rabbit hole, down the rabbit, redpill moment, pill moment, society lies, mainstream lie, controlled narrative

Manosphere / Sexual Strategy

hypergamy, alpha, beta, omega, sigma, delta, gamma, frame, masculinity, masculine, smv, sexual market value, sexual marketplace, tingles, shit test, chad, chadlite, looksmax, looksmaxing, game, dominance, dominant, hierarchy, female nature, branch swing, af bb, alpha fucks, beta bucks, provider, protector, primal, attraction, seduction, pickup, pua, red pill woman, rpw, mgtow, going your own way, monk mode, stoicism, stoic, lift, lifting, gym, testosterone, self improvement, self mastery, discipline, abundance mindset, scarcity mindset, hypergamous, solipsism, female solipsism, dual mating, short term mating, long term mating, alpha widow, cock carousel, wall, the wall, smv decline, post wall, settle down, provider male, beta orbiter, orbiter, simp, white knight, oneitis, pedestalize, pedestalizing, dread game, dread, nexting, spinning plates, plate, plates, outcome independent, amused mastery, dark triad, narcissism, machiavellianism, psychopathy, red pill man, masculine frame, soft harem, open hypergamy, branch swinging, monkey branching

Blackpill / Fatalism

blackpill, black pill, blackpilled, cope, coping, copium, looksism, lookism, subhuman, sub human, sub8, sub6, looksmatch, it is over, its over, over for, noodlewhore, fakecel, truecel, truecels, FA, forever alone, doomed, no point, hopeless, worthless, genetics, height, jawline, canthal tilt, fwahr, facial width, bone structure, bones, skull, nt, neurotypical, autist, autism, social skills, born this way, genetic lottery, genetic trash, low tier, high tier, tier list, mog, mogged, mogging, looksmogged, heightmogged, statusmogged, volcel, incel, inkwell, femoid, foid, foids, stacy, becky, normie, chad, chadfish, chadlite, mentalcel, currycel, ricecel, ethnicel, heightcel, gymcel, moneycel, personalitycel, socialcel, ldar, lay down and rot, rot, rotting, suicide fuel, suifuel, just be, just be tall, just be chad, no hope, no escape, structural fix, fixed hierarchy, fixed outcome, fate, destined, ER, going ER, saint, high score, supreme gentleman, blackpill science, lookspill, heightpill

Counter-discourse / Critical

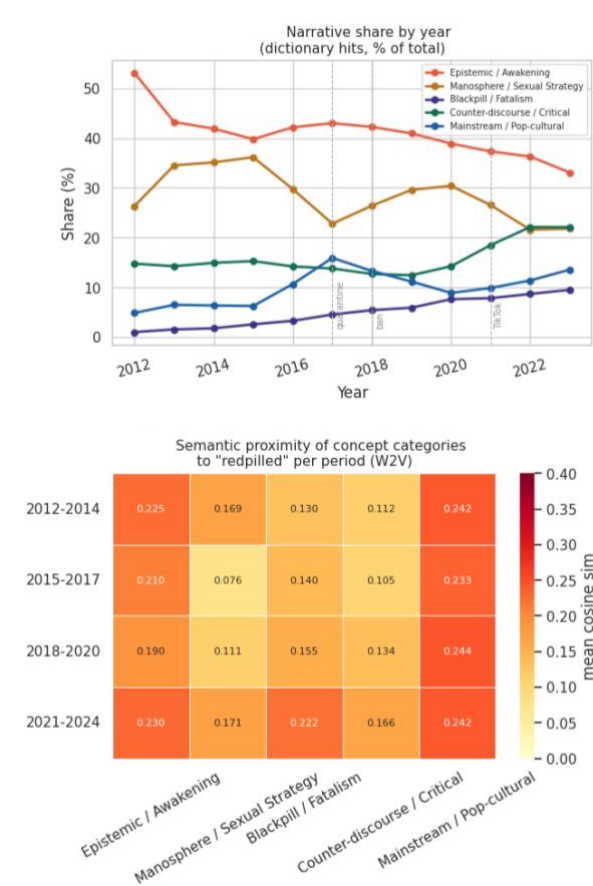
radicalization, radicalisation, pipeline, cult, cult like, dangerous, ideology, misogyny, misogynist, misogynistic, toxic, toxicity, manipulation, manipulative, brainwash, brainwashing, echo chamber, myth, debunk, debunking, criticism, criticize, critique, wake up, leave, escape, deprogram, deprogramming, harmful, harm, hate, hatred, hateful, bigotry, sexism, sexist, pseudoscience, junk science, bad science, misinformation, disinformation, conspiracy, paranoia, paranoid, victim, victimhood, grievance, resentment, incelsphere, manosphere, far right, alt right, extremism, extremist, violent, violence, threat, red flag, warning sign, concerning, worry, worried, ex redpill, former redpill, left the community, peaked, detox, recovery, healthier, therapy, mental health, depression, loneliness, isolation, vulnerable, vulnerability, exploitation, exploited, grooming, predatory, recruitment, gateway, dehumanization, dehumanise, objectification

Mainstream / Pop-cultural

sigma, grindset, hustle, based, cringe, npc, trad, tradwife, trad wife, andrew tate, tate, fresh and fit, podcast, youtube, tiktok, instagram, twitter, social media, motivational, motivation, self help, life coach, entrepreneur, boss, alpha male, alpha mindset, masculine energy, divine masculine, red pill coach, content creator, influencer, clout, views, viral, meme, memes, ratio, based and redpilled, ironically, unironically, actually redpilled, kind of redpilled, lowkey redpilled, ngl redpilled, redpilled about, blackpilled about, everyday feminism, liberal, woke, cancel, cancel culture, political correctness, pc culture, mainstream media,

MSM, fake news, clown world, honkpill, purple pill, pink pill, white pill, doomer, bloomer, coomer, zoomer, boomer, giga chad, gigachad, virgin vs chad, wojak, feels guy, it's giving, slay, queen, king

To characterize the qualitative direction of semantic drift identified through the Word2Vec analysis, a dictionary-based semantic field mapping procedure was employed. Five discrete semantic fields were operationalized as structured keyword dictionaries, with the field categories themselves derived from existing scholarship on manosphere discourse and the sociolinguistics of digital radicalization. Critically, however, the specific lexical content of each dictionary was not imposed from the literature but instead derived empirically from the nearest-neighbor extraction procedure described in the preceding section, ensuring that the vocabulary populating each field reflects the actual distributional patterns of the corpus rather than theoretically presupposed word lists. The five fields are as follows: Epistemic/Awakening, encompassing vocabulary associated with truth-seeking, consciousness, and the Matrix metaphor; Manosphere/Sexual Strategy, comprising the ideological and behavioral lexicon of the men's rights and pickup artist communities; Blackpill/Fatalism, capturing the deterministic and nihilistic vocabulary associated with incel discourse and looks-based hierarchy; Counter-discourse/Critical, indexing the meta-analytical and critical vocabulary through which the term is discussed, challenged, or pathologized from outside the community; and Mainstream/Pop-cultural, reflecting the memeified and increasingly decontextualized usage of the term across broader online culture. Each candidate term extracted from the Word2Vec nearest-neighbor lists was manually assigned to the field whose scholarly definition it most unambiguously satisfied, with terms exhibiting unstable or ambiguous field membership excluded from the final dictionaries. For each calendar year in the study period, all comments in the corpus were scored against the five dictionaries and assigned to the semantic field generating the highest proportion of hits, yielding an annual distribution of field membership across the full 2012 to 2024 timeline. This procedure allowed for the systematic tracking of which semantic frame dominated usage of redpill in any given year, and at what rate the proportional balance among competing frames shifted over the course of the study period.



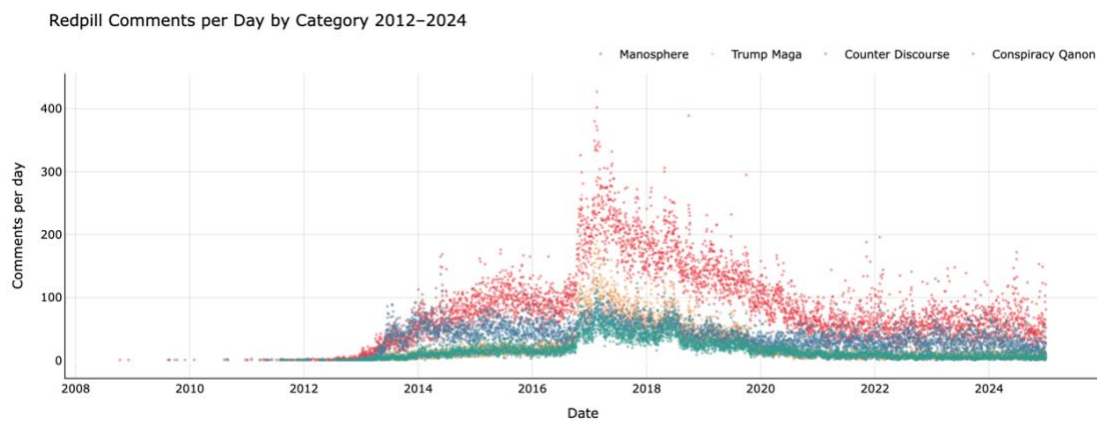
Graph from Dataset 6 *Diachronic Semantic Drift of Redpilled Discourse Across Narrative Categories (2012–2024)*.

Panel 1 shows the narrative share (% of total dictionary hits) of five semantic categories associated with redpilled discourse from 2012 to 2023. The dominant Epistemic/Awakening frame declines gradually over the period, while Mainstream/Pop-cultural and Counter-discourse/Critical uses show consistent growth, indicating progressive semantic broadening. Vertical dashed lines mark the quarantine of r/TheRedPill (2018), a platform intervention coinciding with visible shifts in narrative distribution. Panel 2 presents mean cosine similarity scores (Word2Vec) between redpilled and each concept category across four time periods (2012–2014, 2015–2017, 2018–2020, 2021–2024). Darker cells indicate greater semantic proximity. The consistently high similarity of the Mainstream/Pop-cultural category and the sharp dip then recovery of Blackpill/Fatalism in 2015–2017 reflect the diachronic modulation and cross-community semantic convergence discussed in the analysis.

Semantic Field Classification

To classify the discursive orientation of each comment in the corpus, a semantic similarity classification procedure was employed utilizing the pretrained sentence embedding model all-MiniLM-L6-v2. Four semantic fields were operationalized as natural-language descriptor texts: Origin/Manosphere, capturing the ideological and behavioral lexicon of the men's rights, pickup artist, and incel communities; Trump/MAGA, encompassing the populist right, anti-establishment, and nationalist political vocabulary with which the term became increasingly associated from 2016 onward; Counter-discourse/Critical, indexing the meta-

analytical, ironic, and critical vocabulary through which the term is discussed or pathologized from outside the community; and Conspiracy/QAnon, reflecting the epistemically paranoid and hidden-truth framing characteristic of broader online conspiracy culture. Each category was represented as a dense semantic description text, which was encoded into a fixed-dimensional embedding vector using all-MiniLM-L6-v2. All 780,000 comments in the corpus were similarly encoded in batches, with checkpointing implemented to ensure continuity across processing sessions. Each comment was then assigned to the semantic field whose descriptor embedding exhibited the highest cosine similarity to the comment embedding, with similarity scores for all four fields retained in the output to enable confidence assessment and longitudinal aggregation. Field membership was subsequently traced across each calendar year from 2012 to 2024, yielding an annual distribution of semantic field proportions that documents the shifting discursive contexts within which *redpill* was deployed across the study period.



Graph from Dataset 7 Redpill Comments per Day by Category, 2012-2024

Daily volume of redpill-mentioning comments across four semantic fields (Manosphere, Trump/MAGA, Counter-discourse, Conspiracy/QAnon), classified via cosine similarity to field descriptor embeddings using all-MiniLM-L6-v2. Each dot represents one day's comment count within a given field. The corpus remains almost exclusively Manosphere-coded until 2013, after which gradual growth across all fields is visible. A sharp volumetric peak emerges around 2016-2018, coinciding with the political mainstreaming of redpill rhetoric and the rise of Trump/MAGA discourse as a dominant semantic field. Post-2018, overall volume declines but stabilises at a higher baseline than the pre-2016 period, with Manosphere and Counter-discourse comments persisting as the two most consistent fields through 2024.

(3) Community Semantic Profile (CSP)

For a given subreddit s , let $F = \{f_1, f_2, \dots, f_n\}$ be the set of semantic fields. The semantic profile of s is defined as:

$$CSP(s, f_i) = \frac{\sum_{c \in s} \mathbb{1}[\hat{f}(c) = f_i]}{|s|}$$

C = comment

$\hat{f}(c)$ = assigned field for comment c

|s| = total number of comments in subreddit s

The calculation of the CSP allows to plot each subreddit within a barycentric position, returning 3 major hubs of semantic use.



Graph from Dataset 8 Semantic Space of Subreddit Discourse

Triangular semantic map situating all redpill-associated subreddits according to the relative distribution of semantic field classifications within their comment corpora, as produced by the all-MiniLM-L6-v2 classification procedure. Each bubble represents one subreddit; bubble size encodes total comment volume. Position is determined by barycentric interpolation between three poles - Manosphere (top-left), Far Right aka. MAGA/Conspiracy (top-right, with Conspiracy/QAnon merged into the MAGA field), and Counter-discourse (bottom) - such that the geometric center corresponds to an equal 33% distribution across fields. The map reveals a clear structural divide: the majority of manosphere-adjacent communities cluster in the upper-left quadrant, with explicitly political and conspiracy-adjacent subreddits (r/The_Donald, r/DebateAltRight, r/greatawakening) occupying the MAGA/Conspiracy pole, while counter-discursive communities (r/exredpill, r/TheBluePill, r/niceguys) congregate near the bottom vertex

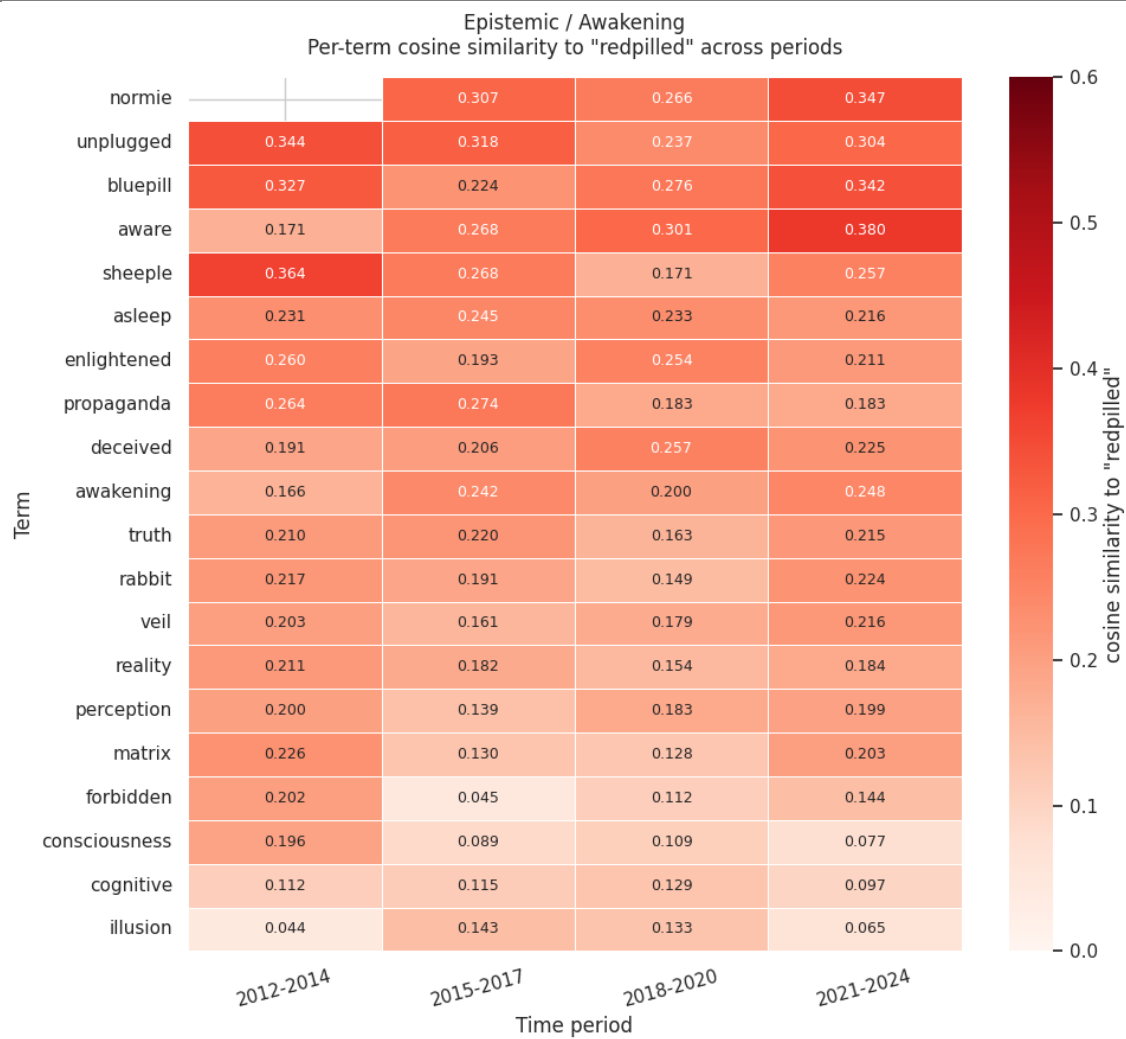
The methodological procedures outlined above yield a multi-layered quantitative foundation from which the substantive analysis of redpill's semantic trajectory can proceed. Taken together, the Word2Vec drift measures, dictionary-based narrative tracking, sentence-embedding classification, and community semantic profiling converge on a coherent preliminary picture. The term originates as a densely coded manosphere-internal lexical item,

almost exclusively Epistemic/Awakening in register and tightly bound to the sexual strategy and truth-seeking vocabulary of *r/TheRedPill's* founding community. From approximately 2013 onward, and with marked acceleration around 2015-2016, the term undergoes rapid cumulative semantic displacement, departing substantially from its baseline vector and acquiring distributional proximity to political, conspiratorial, and eventually mainstream pop-cultural vocabulary. Three discrete moments of heightened incremental drift in 2013, 2016, and 2020 locate the precise temporal coordinates of this lexical rupture, each corresponding to documented events in the broader socio-political and platform environment: initial community diffusion, the politicization of the term during the Trump campaign and alt-right consolidation, and the convergence of pandemic discourse with QAnon mainstreaming during the 2020 election cycle. Post-2020, drift stabilizes, suggesting the term has reached a semantically diversified but relatively settled distributional equilibrium. At the community level, the CSP mapping reveals a structural tripartition of redpill discourse across subreddits: a manosphere-dominant cluster, a far-right and conspiracy-adjacent cluster, and a counter-discursive cluster, demonstrating that the term does not occupy a single semantic space but simultaneously sustains distinct and partially incompatible discursive functions depending on the community deploying it. Also, the content moderation informed shut-down and quarantining of certain Subreddits is reflected in this drift line; this study finds that some communities serve as specific semantic hubs, incubating partly distinct definitions of the term, when these disappear so does this definition prominence

Results¹⁷

Proximity of concept words over time periods

1 Proximity of concept words from the Epistemic/Awakening dictionary to "redpilled" over four time periods (2012–2024)

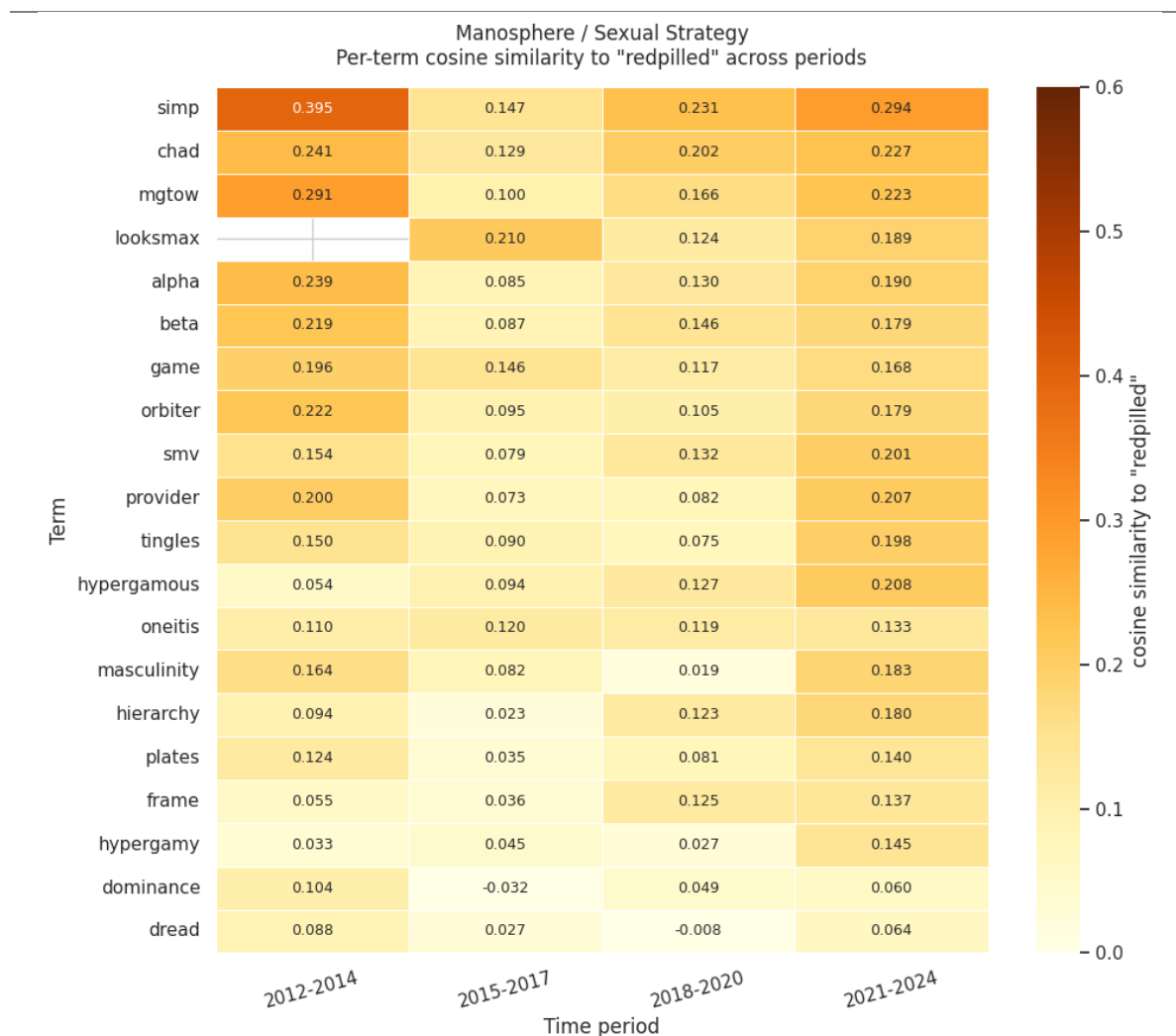


Notable patterns: include the rising similarity of "aware" (0.171 -> 0.380) and "normie" (n/a -> 0.347) or "sheeple" (0.364 -> 0.257), suggesting a shift from philosophical framing toward in-group/out-group identity language, while originally strong collocates like and "matrix" (0.226 -> 0.203) weakened as the term departed from its pop cultural origins.

2 Proximity of concept words from the Manosphere dictionary to "redpilled" over four time periods (2012–2024)

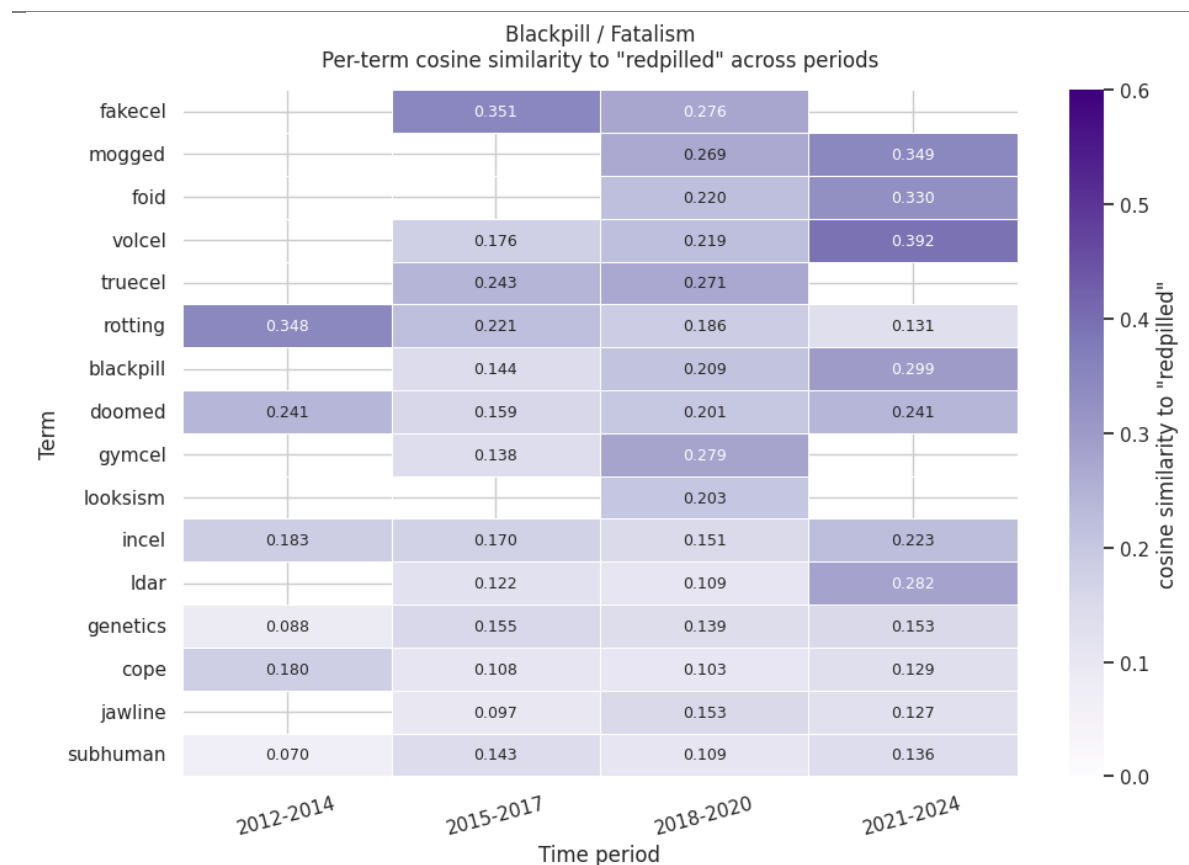
¹⁷ A summary of all results is available under following link:

Due to the word limitation of thesis some results will just be sketched in a table or linked in the Appendix



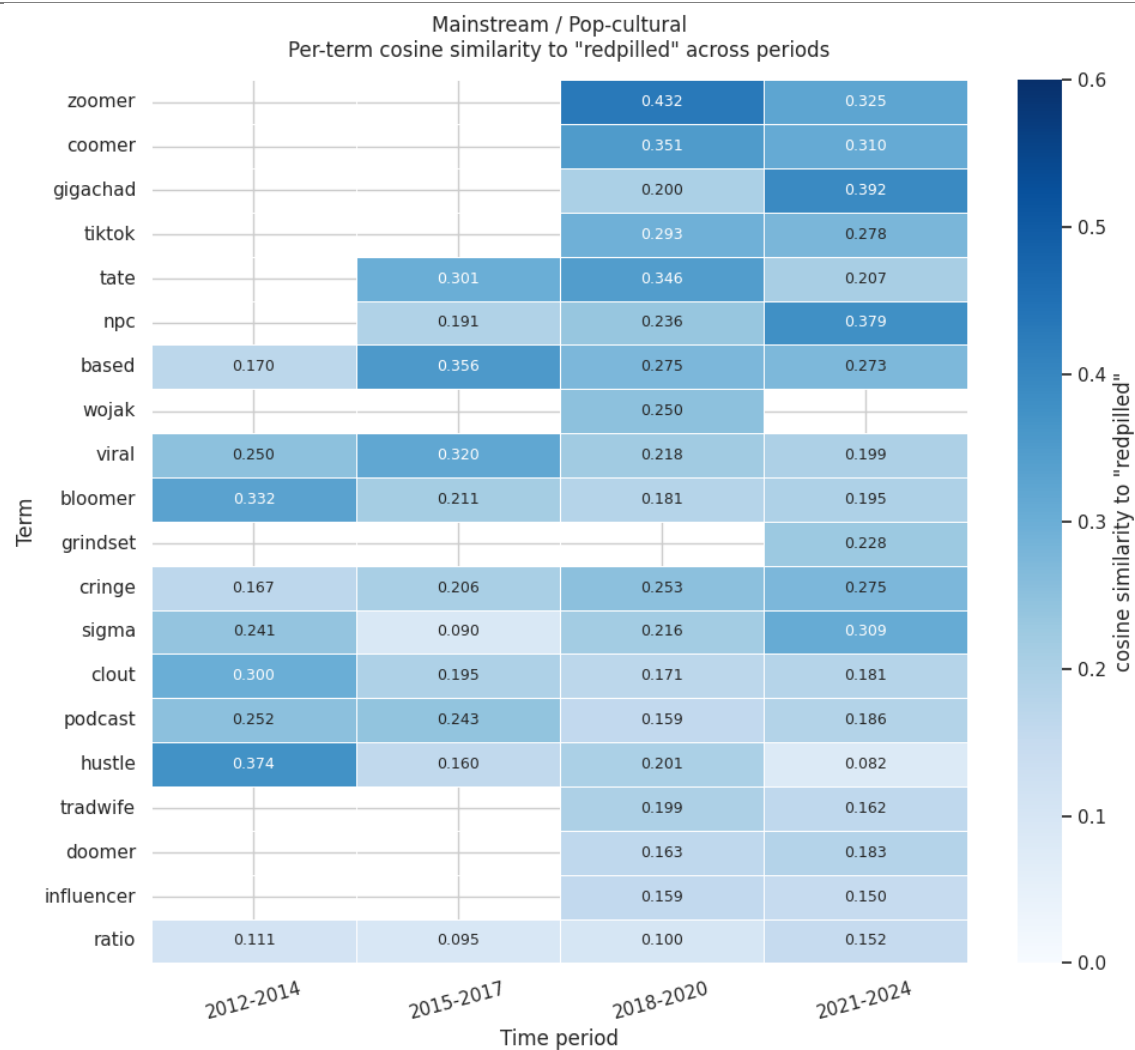
Notable patterns: Terms peak in the early manosphere era (2012-2014), dip during the political migration of 2015-2017, then rebound by 2021-2024, particularly for terms that entered the mainstream through influencer culture (Bujdei-Tebeica, 2023). "*Hypergamous*" (0.054 -> 0.208), "*hierarchy*" (0.094 -> 0.180), and "*hypergamy*" (0.033 -> 0.145) all roughly tripled in similarity, indicating that pseudo-evolutionary vocabulary became more tightly bound to *redpilled*. The early-period dominance of "*simp*" (0.395) weakened mid-decade before partially recovering. This suggests that the manosphere lexicon temporarily lost associative ground to sexual language during 2015-2017 before reasserting itself after. This can also be linked to the supposed deliberative phase towards the end of the dataset where the term's meaning might have been discussed in a meta-analytical and retrospective fashion by the outside group.

3 Proximity of concept words from the Blackpill dictionary to "*redpilled*" over four time periods (2012–2024)



Notable patterns: The sharpest trajectories belong to "*volcel*" (0.176 -> 0.392), "*mogged*" (n/a -> 0.349), and "*foid*" (n/a -> 0.330), all of which appear only in later periods and climb steeply, marking the crystallization of incel-specific language and its growing semantic proximity to *redpilled*. "*Ldar*" (lie down and rot) surges to 0.282 by 2021-2024, and "*blackpill*" itself rises from 0.144 to 0.299, suggesting an increasing simultaneous use of the red-pill and black-pill semantic fields, a computational reflection of the radicalization pipeline documented in qualitative research. The early outlier "*rotting*" (0.348 in 2012-2014) collapses across later periods (0.131), possibly reflecting its absorption into the more codified term "*ldar*." This category is distinguished by the late emergence of most terms, as many cells are empty in 2012-2014, reflecting a vocabulary that largely did not yet exist.

4 Proximity of concept words from the Mainstream dictionary to "*redpilled*" over four time periods (2012–2024)



Notable patterns: "Zoomer" (n/a -> 0.432 -> 0.325), "coomer" (n/a -> 0.351), "gigachad" (n/a -> 0.392), and "tiktok" (n/a -> 0.278) all materialize in 2018-2020 with no prior presence, reflecting TikTok-era virality and the Gen Z audience. "Sigma" nearly triples (0.090 -> 0.309) and "npc" climbs steadily (0.191 -> 0.379), signaling the consolidation of internet-native identity archetypes around *redpilled*. Conversely, earlier culture-war markers like "hustle" (0.374 -> 0.082) and "clout" (0.300 -> 0.181) weaken. The appearance of "tradwife" (n/a -> 0.199 -> 0.162) and "grindset" (n/a -> 0.228) only in later periods marks the term's expansion into potential lifestyle branding.

Discussion

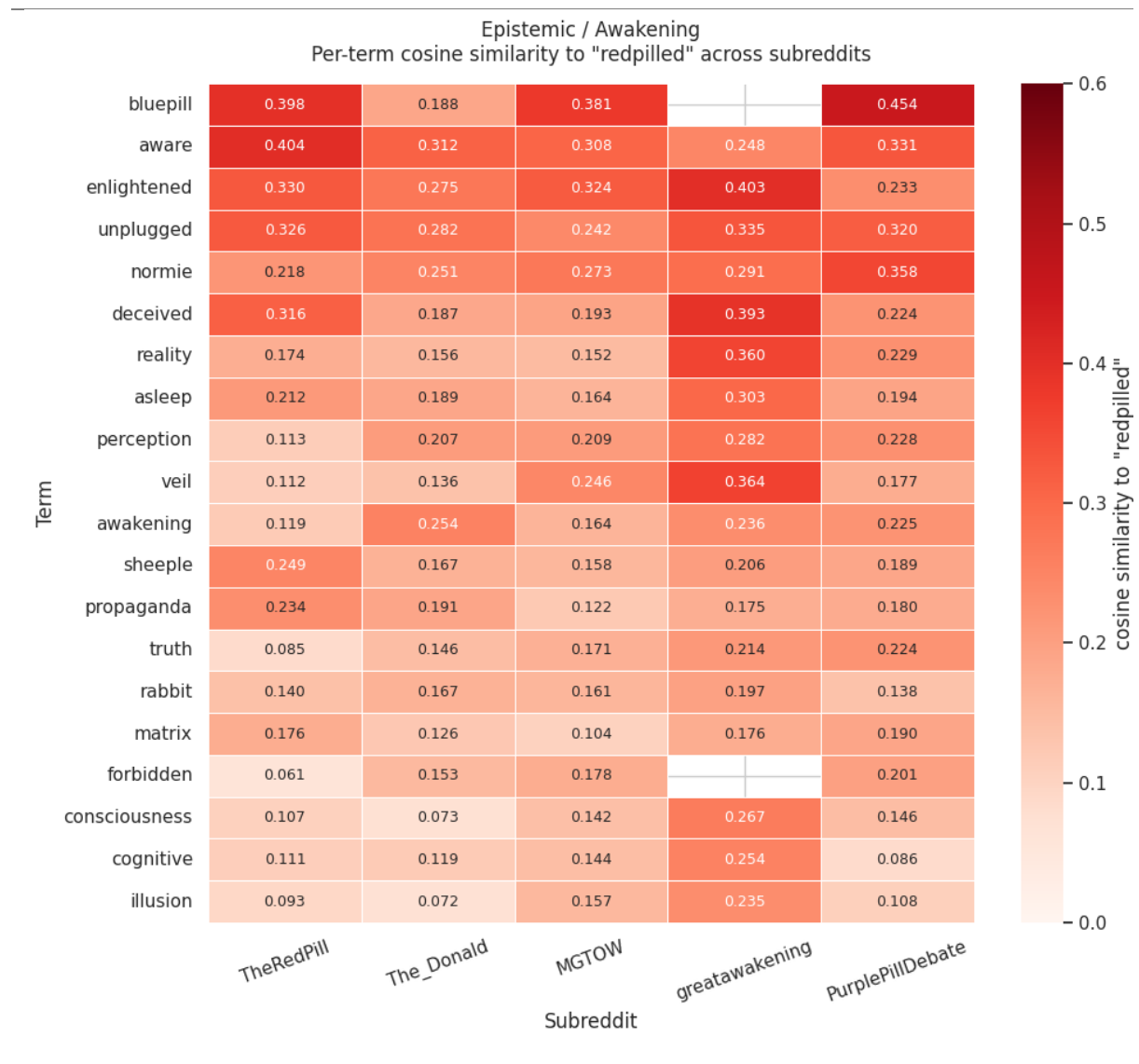
First, a temporal U-curve is visible across the Epistemic and Manosphere categories: terms strongly associated with *redpilled* in 2012-2014 weakened during 2015-2017, the period of political migration through Gamergate and the Trump campaign, before rebounding by 2021-2024. This suggests political mainstreaming temporarily diluted the term's subcultural specificity, but the original semantic domains re-consolidated once the term stabilized in broader usage.

Second, the Blackpill and Mainstream categories show a late-emergence pattern: most of their strongest terms simply did not exist before 2015 or 2018. The vocabulary surrounding *redpilled* is not merely shifting, it is actively generating new lexicon. Terms like "*looksmax*," "*foid*," "*gigachad*," "*grindset*," and "*npc*" were coined within the communities that use *redpilled* and have become increasingly bound to it.

Third, the data tracks a shift from epistemological framing to identity taxonomy. In 2012-2014, *redpilled* was most strongly associated with knowledge-state terms (*sheeple*, *unplugged*, *bluepill*, *matrix*), language about what one knows. By 2021-2024, the highest-similarity terms across all four categories are increasingly identity markers (*normie*, *aware*, *npc*, *sigma*, *gigachad*, *volcel*), language about what one is. This reflects the broader trajectory: *redpilled* has evolved from describing an epistemological event (a moment of awakening) to encoding a social identity (membership in a community of the awake).

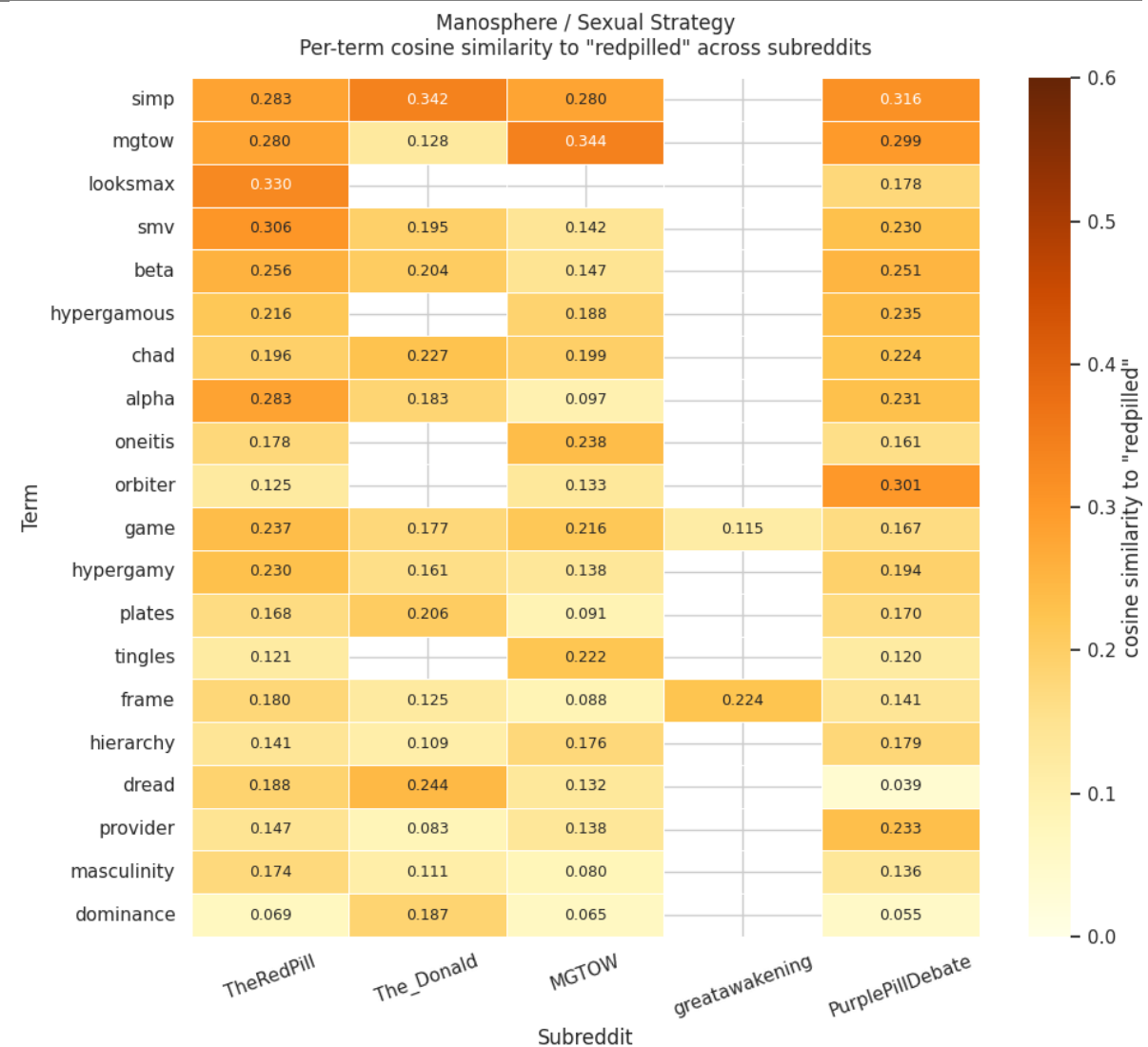
Proximity of concept words over Subreddits

1 Proximity of concept words from the Epistemic dictionary to "*redpilled*" over four biggest subreddits (2012–2024)



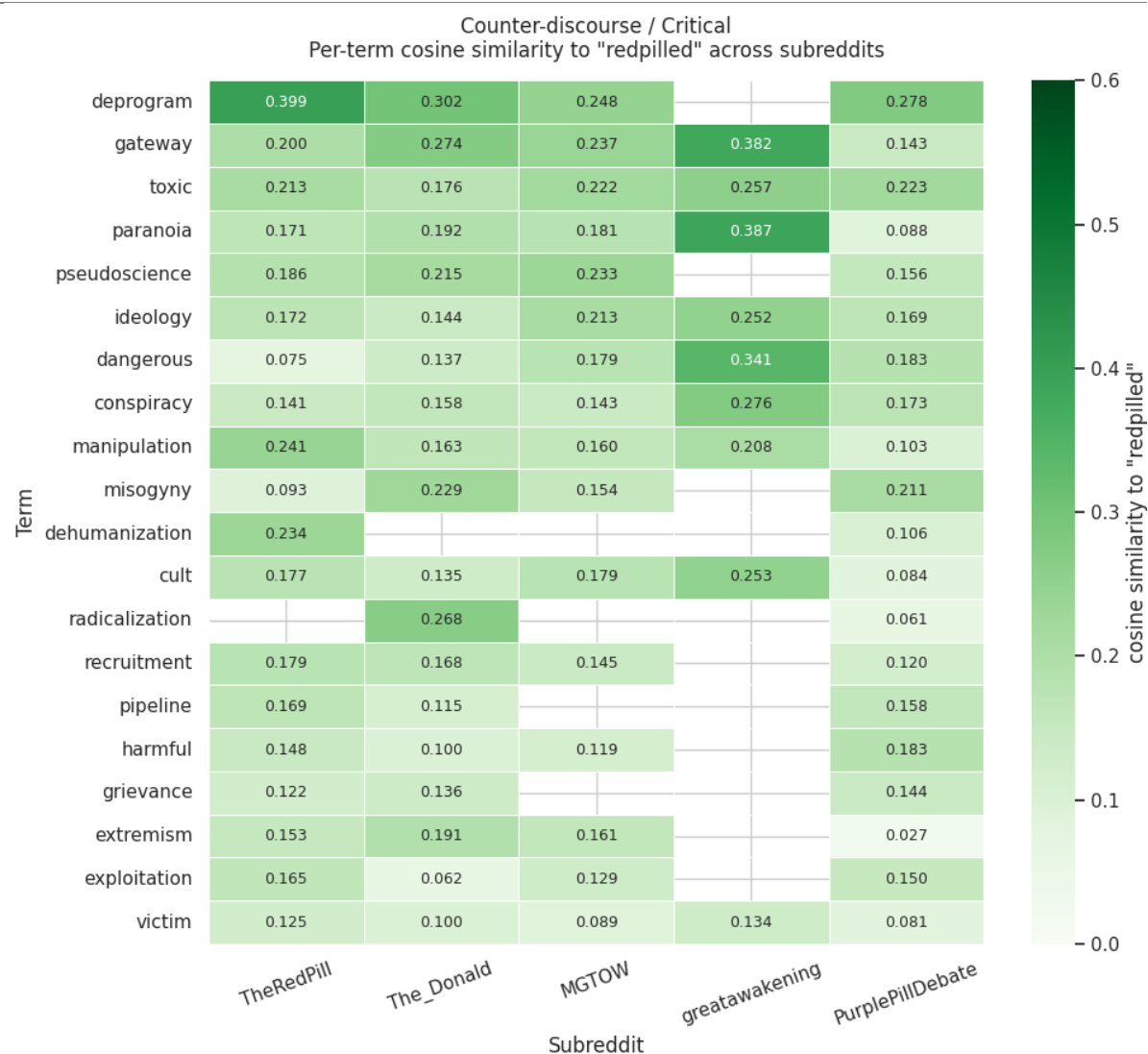
Notable findings: *r/greatawakening* dominates the epistemic register, producing the highest similarity scores for "*deceived*" (0.393), "*veil*" (0.364), "*reality*" (0.360), and "*enlightened*" (0.403), suggesting that QAnon discourse embedded *redpilled* most deeply within a framework of hidden truths and concealed realities. By contrast, *r/TheRedPill* and *r/PurplePillDebate* cluster most strongly around "*bluepill*" (0.398 and 0.454 respectively) and "*aware*" (0.404 and 0.331), indicating that the manosphere uses epistemic language primarily to draw binary in-group/out-group distinctions rather than to elaborate conspiracist narratives. *r/The_Donald* sits in between, with moderate similarity across all terms and no single dominant cluster, reflecting its role as a generalist political community that borrowed awakening language without anchoring it to a specific epistemic system. "*Sheeple*" (0.249) and "*propaganda*" (0.234) load most heavily on *r/TheRedPill*, suggesting that early manosphere usage retained more of the original Matrix-era contempt for the unawakened than later political adoptions did.

2 Proximity of concept words from the Manosphere dictionary to "*redpilled*" over four biggest subreddits (2012–2024)



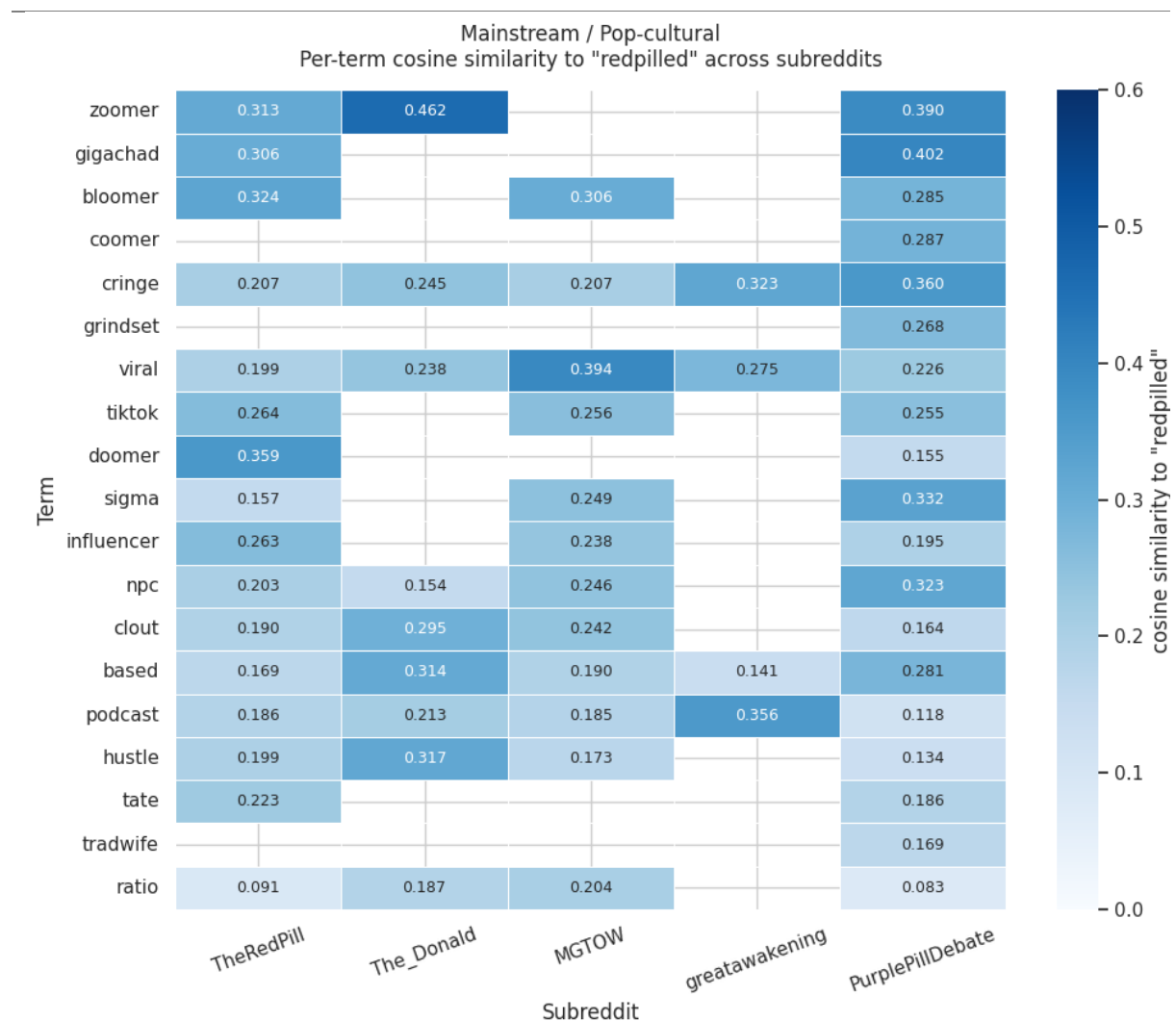
Notable findings: The manosphere lexicon splits sharply along community lines. "Simp" is the only term with high similarity across all three manosphere-adjacent subreddits (*TheRedPill* 0.283, *The_Donald* 0.342, *MGTOW* 0.280), confirming its crossover status as a term that migrated from dating culture into political insult. *r/TheRedPill* predictably dominates terms of sexual marketplace theory: "*looksmax*" (0.330), "*smv*" (0.306), "*alpha*" (0.283), and "*hypergamy*" (0.230) all peak there. *r/MGTOW* loads most heavily on "*mgtow*" itself (0.344) and "*tingles*" (0.222), reflecting its more withdrawal-oriented stance. *r/The_Donald's* strongest manosphere term is "*dread*" (0.244), a manipulation tactic from pickup artistry, suggesting selective borrowing of the most adversarial relationship concepts into political discourse. *r/greatawakening* is almost entirely absent from this category, with most cells empty, confirming that QAnon's appropriation of *redpilled* operated independently of the gendered manosphere lexicon. *r/PurplePillDebate* shows elevated scores for "*orbiter*" (0.301), "*provider*" (0.233), and "*beta*" (0.251), consistent with its function as a debate forum.

3 Proximity of concept words from the Critical dictionary to "*redpilled*" over four biggest subreddits (2012–2024); not featured in results I



No clear patterns could be established.

4 Proximity of concept words from the Mainstream dictionary to "redpilled" over four biggest subreddits (2012–2024)



Notable findings: "Zoomer" peaks dramatically in *r/The_Donald* (0.462), the highest single value across all four heatmaps, suggesting that political communities were the primary site where generational identity language fused with *redpilled*. "Gigachad" concentrates in *r/PurplePillDebate* (0.402) and *r/TheRedPill* (0.306) rather than political subreddits, confirming its manosphere origins before broader adoption. "Based" loads most heavily on *r/The_Donald* (0.314) and *r/PurplePillDebate* (0.281), consistent with its function as a political affirmation term (Important caveat: "*based*" cannot be isolated clearly from the passive voice to be based). "Viral" peaks in *r/MGTOW* (0.394), and "podcast" peaks in *r/greatawakening* (0.356), reflecting each community's distinct media consumption patterns. "Tradwife" and "grindset" appear almost exclusively in *r/PurplePillDebate* (0.169 and 0.268), suggesting these lifestyle-branding terms entered red-pill discourse through the debate community rather than the ideological core. "Cringe" shows its highest similarity in *r/PurplePillDebate* (0.360) and *r/greatawakening* (0.323), functioning as a social policing term across both gender and political registers.

Discussion

The subreddit-level analysis reveals that *redpilled* does not carry a single meaning that varies in intensity across communities. Instead, each subreddit activates a distinct semantic profile, effectively making *redpilled* a different word in each space.

r/greatawakening operates *redpilled* almost entirely within the Epistemic and Counter-discourse categories, with near-total absence from the Manosphere lexicon. This confirms that QAnon's *redpilling* was a parallel appropriation that shared the metaphor's epistemological structure (*hidden truth, mass deception, awakening*). The conspiratorial register (*paranoia, dangerous, conspiracy*) is unique to this community.

r/TheRedPill and *r/MGTOW* share a manosphere core but diverge somewhat in posture. *TheRedPill* loads on terms of sexual competition and self-improvement (*looksmax, smv, alpha*), while *MGTOW* loads on withdrawal and contempt (*mgtow, tingles, viral*). Both are largely absent from mainstream pop-cultural terms, thus these communities may have remained more insular even as their vocabulary leaked outward.

r/The_Donald functions as a mixing chamber. It borrows selectively from the manosphere (*simp, dread, chad*), absorbs mainstream internet culture (*zoomer, based, hustle, clout*), and registers the highest exposure to critical discourse about political radicalization. Its semantic profile for *redpilled* is the broadest and most diffuse, consistent with its role as a large-tent political community where multiple red-pill meanings coexisted.

r/PurplePillDebate is the most metacognitive space. It scores consistently on terms from every category without dominating any single one, and it uniquely concentrates lifestyle-branding terms (*tradwife, grindset, gigachad, cringe*) alongside critical vocabulary (*harmful, exploitation, misogyny*). This positions it as the community most likely to use *redpilled* reflexively, discussing the concept rather than performing it. Thus, despite sharing the same metaphor, these communities developed functionally independent semantic ecosystems around *redpilled*, converging only on the broadest epistemic terms (*aware, enlightened, deceived*). This suggests that the red-pill metaphor lies not in any fixed content but in its empty formal structure (Cf. Lexicographic Profile). a binary choice between comfortable ignorance and painful truth, which any community can fill with its own specific 'truth'.

At this analytical junction it remains indispensable to recall the findings of figure 17¹⁸ that portray clear incubatory spaces around the major Subreddits and their semantic uses. These communities center around larger bubbles whose deployment of the term are often mirrored in

¹⁸ (semantic fields have been cleaned since MAGA/Conspiracy largely overlapped)

other communities. It can be speculated that a mix of specific term culture within a subreddit and the attainment of a critical mass of comments drives semantic change. However, most semantic incubations are often stopped before reaching critical mass (ex. decline of conspiratorial vocabulary incubated by *r/thegreatawakening*) due to content moderation events. If they succeed to reach semantic ingenuity and critical mass they usually become semantic drivers. The study identifies *r/The_Donald*, *r/PurplePillDebate*, *r/TheRedPill*, *r/MGTOW*, *r/TheBluePill*, and *r/exredpill*, *r/thegreatawakening*.

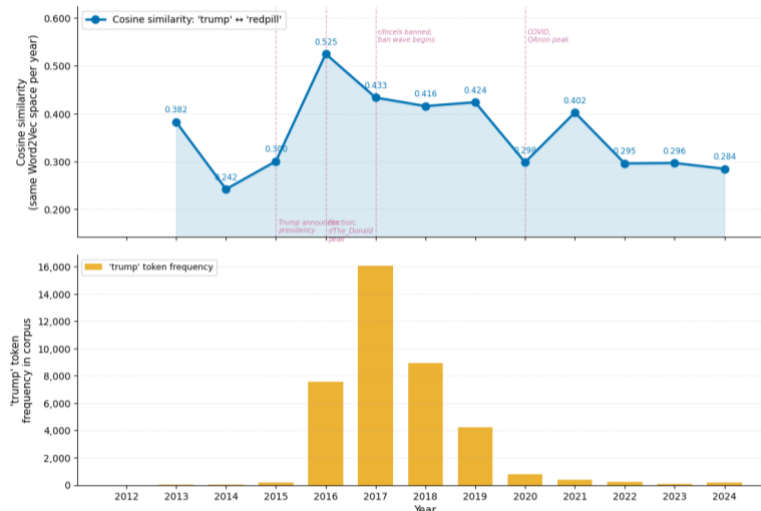


Figure 18 Semantic Space of Subreddit Discourse

Take *r/The_Donald* as an example that started mostly by reappropriating manosphere language from *r/theredpill*, linking it to the 2016 presidential election and thus the figure of Donald Trump as *redpilled*. As a passionate support *u/jesseale* reaffirms in *r/The_Donald*:

“Trump support is surging. Expect 60% by Tuesday. GO VOTE! WE NEED A RECORD BREAKING LANDSLIDE TO OVERCOME THIS RIGGED ELECTION. START RED PILLING YOUR FRIENDS AND FAMILY ON THE HORRORS OF HILLARY CLINTON”.

Soon specific actors, which suspicious names (ex. *___TRUMP2016___*) started using the term in connection with the Red pill conservative journalist, starting with said subreddit. The analysis could not surely establish whether these actors are bots or real Redditors, suggesting a need for further clarification. Still, the rise and fall of Trump as a semantic neighbor, whether strategic or naturally grown, can equally be traced (cf. Graph from Dataset 9).



Graph from Dataset 9 Semantic Proximity of 'trump' to 'redpilled', Year by Year
Correlation Matrix

Following correlations could be retrieved from the data set (divided into two periods, diffusion and consolidation phase). These are especially present in the early years of the data set before content moderation events started to externally influence the described linguistic communities. The study places no judgment on the ethical and political dimension of these decisions yet highlights their evident influence on the term’s semantic trajectory. It is understood that the specific makeup of constraint shaped registers finetunes semantic evolution (Cf. CSR). Significant correlations (p; **) exist between drift and volume as well as drift and diversity of actors (Subreddit and Redditor diversity). Strongly significant correlations (p; ***) exist between the increasing use of the word and its diversity of actors. It may be deduced that the more the word *redpilled* is used the more diverse its semantic profile becomes; and the more diverse the actors using it. This is an empirical proof to this otherwise seemingly obvious hypothesis (H1). The study acknowledges as well that more accurate analytical tools and a better cleansed data set would naturally produce stronger correlations. Especially that some actors could not have been identified as either bots or fanatics weaksn the present analytical deductions,

Period 1: 2013–2018 (n = 6 years; diffusion phase)					
Pair	Pearson r				p
Drift × Comment count	+0.819	0.046	*	+0.943	0.005 **
Drift × Subreddit diversity	+0.931	0.007	**	+0.943	0.005 **
Drift × Redditor diversity	+0.838	0.037	*	+0.943	0.005 **
Comment count × Subreddit div.	+0.954	0.003	**	+1.000	0.000 ***

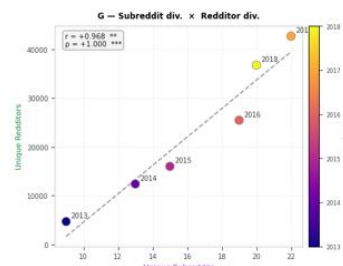
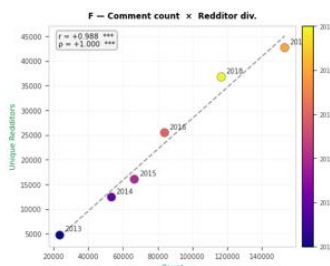
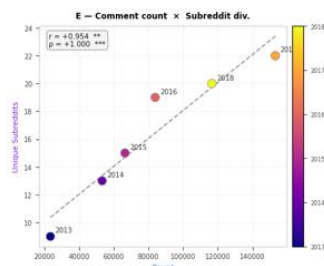
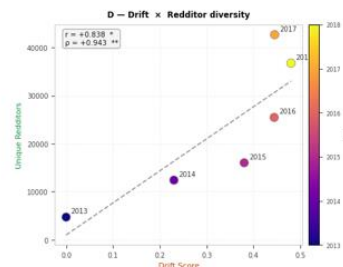
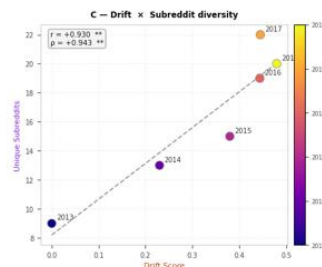
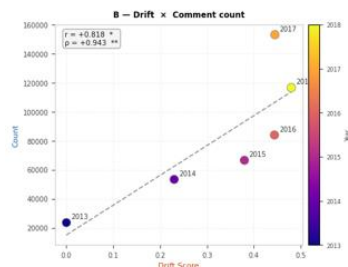
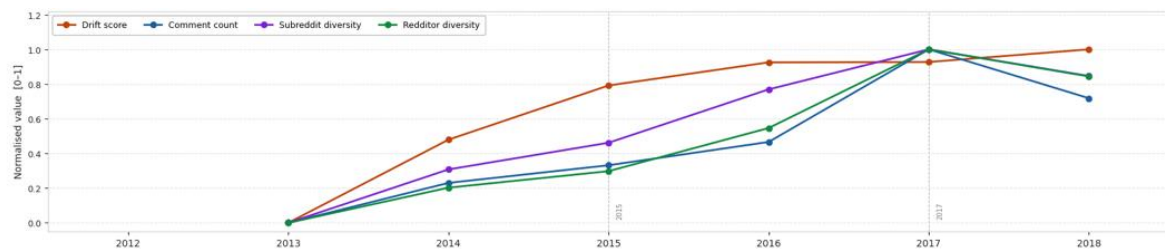
Comment count × Redditor div.	+0.988	0.000	+1.000	0.000
		***		***
Subreddit div. × Redditor div.	+0.968	0.002	+1.000	0.000
		**		***

Period 2: 2019–2024 (n = 6 years; consolidation phase)

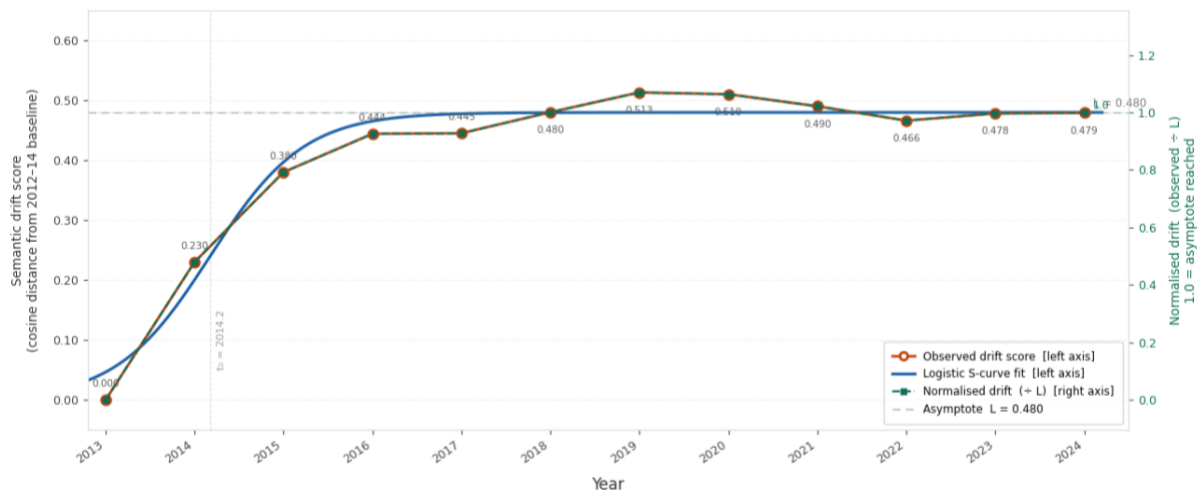
Pair	Pearson r		p	
Drift × Comment count	+0.787	0.063 n.s.	+0.829	0.042 *
Drift × Subreddit diversity	+0.952	0.003 **	+0.926	0.008 **
Drift × Redditor diversity	+0.827	0.042 *	+0.771	0.072 n.s.
Comment count × Subreddit div.	+0.818	0.047 *	+0.926	0.008 **
Comment count × Redditor div.	+0.987	0.000	+0.943	0.005 **

Subreddit div. × Redditor div.	+0.865	0.026 *	+0.926	0.008 **

Significance: *** $p < .001$ ** $p < .01$ * $p < .05$ n.s. not significant · Pearson r = linear · Spearman ρ = rank-order



Graph from Dataset 10 "Redpilled" Correlation Analysis Semantic Drift x Volume x Community Diversity n = first 6 year-observation (2012-2018) Pearson r + Spearman ρ ; after content. Moderation events (year=2018) correlations weaken since use of the term is externally mediated



Graph from Dataset 11 “Redpill” Semantic drift score to logistics S-curve, Asymptote $L = 0.480$, interpreted as core semantic root that towards infinity cannot be semantically omitted, unless unexected external change might occur going against the trends and actors captured in this dataset

Limitations

The study is subject to several limitations that qualify the scope of its findings. At the level of data integrity, the Pushshift.io archive contains the aforementioned constitutive gaps. The approximately 182,000 removed entries cluster disproportionately within high moderation communities during their most ideologically active periods, meaning the corpus likely understates both the volume and discursive intensity of precisely those spaces where the term's semantic charge was most concentrated. This is intrinsic to the archival infrastructure and not correctable post hoc. At the level of actor identification, the pipeline could not reliably distinguish bots and coordinated inauthentic accounts from organic users, leaving open whether certain observed semantic shifts were strategically manufactured or naturally emergent and weakening the correlational findings around drift and actor diversity. While most obvious bots were removed, some could not be excluded definitively. Methodologically, Word2Vec models trained on annual slices remain sensitive to volume thresholds, and the dictionary classification, while empirically derived from nearest neighbor extraction, still requires manual assignment of terms to semantic fields, introducing a degree of interpretive subjectivity. More fundamentally, distributional semantics captures co-occurrence patterns but not meaning. Without mixed method triangulation through close reading, discourse analysis, or ethnographic engagement with community norms, cosine distances and nearest neighbor shifts risk overinterpretation. A vector moving closer to conspiratorial vocabulary does not on its own confirm that users intended conspiratorial meaning; only qualitative contextualization can distinguish genuine semantic reappropriation from incidental co-occurrence. Finally, the

analysis is confined to one platform and one term family, and the post 2020 decline in comment volume means the consolidation phase rests on thinner empirical ground. Content moderation events function as exogenous shocks that externally mediate the linguistic processes under study, and the marked weakening of correlations after 2018 makes it difficult to disentangle organic semantic evolution from platform imposed disruption.

Still, this quantitative groundwork critically furnishes the empirical basis from which a working base definition and etymological profile of the term can be responsibly constructed. This analysis is grounded in the actual distributional history of the word across its full arc of historic use cases. It commences with a data-based lexicographic profile and semantic field entry, subsequently cross-referenced against existing qualitative literature to verify whether the empirically derived definitions are reflected in and supported by prior scholarly work.

Final Results

Lexicographic Profile and Semantic Field Entry

Etymology

Headword:

redpill (also red pill, red-pill; attested forms include redpilled, redpilling, redpillier, redpillers, pre-redpill, de-redpill, redpill as verb, noun, and adjective; see below for full morphological inventory)

Working Definition:

Redpill(ed) is a digitally-native ideological and political lexical token, originating from the cinematographic metaphor in *The Matrix* (Wachowski & Wachowski, 1999), and subsequently reappropriated by online masculinist communities to denote an "awakening" to the alleged systemic privileging of women and the decline of male power, promoting evolutionary determinism and male dominance ([Özbilgin & Mergen, 2025, p. 4](#)). In a broader literary register, being redpilled represents conversion or entryway into a new ideological realm, hence a form of enlightenment or rite of passage through which an individual recognizes an oppressive structure and overcomes it through that very recognition. Therefore, it also includes a general opposition against myriad dominant cultural values that undergird the project of liberal democracy, particularly with respect to gender and racial equality ([Gray & Bernatzky, 2022](#)). As a digitally native lexical item, *redpilled* is notable for its cross-community range, and therein its capacity for measurable semantic expansion over time, and to function simultaneously as sincere ideological commitment, ironic commentary,

and post-ironic cultural gesture. This study terms these forms plurosemantic.

Antonym:

To be woke (from to awake): conscious of systems of oppression, identity politics, or political injustice, especially related to race, sexuality and gender; underpinned by political projects such as racial justice, feminism, or equality advocacy. Metaphor: waking up to the world's injustices; Also used pejoratively as an abstinent identity-based ideology, or "un-redpilled"

Origins:

Shows the evolution of etymological definitions as described in the literature and within sample comments of the data set; remains mostly anecdotal.

- (1) In the *The Matrix* (1999) [Film] the protagonist, Thomas Anderson, known as the hacker Neo (Keanu Reeves), experiences strange occurrences in his life. He is soon contacted by underground figure Morpheus and Trinity, through who he learns that his world is a sophisticated digital simulation. He may return to his old life by taking the blue pill or help humanity escape their coded prison by taking the red pill. The choice is between comfortable bliss or the unlivable truth.
 - (1.1) Draws clear parallels to classic awakening motifs: platonic cave, the rabbit hole, or New Age awakening
 - (2) First political use case by Kathleen J. Tierney after Hurricane Kathrina, saying those shocked by the governments poor response "should take the red pill" and realize that "*post-September 11 policies and plans have actually made the nation more vulnerable, both to natural disasters and to future terrorist attacks*" ([Tierney](#), 2006). In 2007 neo reactionary blogger and software engineer Curtis Yarvin (Mencius Moldbug), whose philosophy is especially present in the second Trump administration and who advocates for an anti-liberal political project, used the term the first time in a clear ideological sense. In the essay titled "*The Case Against Democracy: Ten Red Pills*" Yarvin theorizes the basis of his anti-democratic ideology. Alludes to a great awakening.
 - (3) Since 2012 the term is used primarily by *manosphere* thinkers (see below for full comments from the data set). Emerging as a meme within the r/TheRedPill subreddit, its manosphere definition was coined by Robert Fisher. It largely aimed at promoting traditional forms of masculinity and opposition to feminism (unclearly defined). Frames men as victims of feminism and political correctness. Race and gender were
-

only later included. Associated terms are often hypergamy, alpha/ beta, feminist, sexual marketplace, masculinity, cuck, Chad. Agentive as it recommends young men to fight this regime. Later peripherally linked to the Donald Trump's first presidential campaign.

- (4) Ideological Expansion and Mainstreaming: From 2015 onwards, redpilled undergoes a process of rapid ideological expansion and progressive mainstreaming that fundamentally transforms its semantic range. Beginning with the [GamerGate](#) ([Massanari](#), 2017) controversy of 2014–2015, the term migrates out of its manosphere origin community into a broader culture-war register, where awakening to feminist influence in media and journalism replaces sexual-strategy awareness as its primary referent. The alt-right subsequently loads the term with racial content, while preserving its epistemological template intact, producing the first layer of semantic stratification. Donald Trump's 2016 electoral campaign and victory mark the term's peak political saturation (the corpus data set records all time high mentions in the r/The_Donald alone in February 2017); new meaning: to redpill someone becomes a transitive act of political conversion. The QAnon ecosystem from 2017 re-literalizes the Matrix metaphor, inflating the term into a millenarian register. Simultaneously, mainstream conservative influencers such as Candace Owens, Joe Rogan, Nick Fuentes, and later Andrew Tate and other decouple the term from its subcultural origin, normalizing it as a generic marker of political or ideological conversion for mass audiences with little reference to its manosphere genealogy. By 2021, semantic bleaching is well underway: the -pilled suffix detaches entirely from its ideological charge and circulates ironically as a content-neutral enthusiast marker across mainstream social media. The result is a term that by the mid-2020s simultaneously functions as a live ideological instrument in its origin communities, a mainstream political shorthand, a conspiracist signal, and a content-evacuated ironic suffix. An extreme semantic derailment into myriad of far right use cases has been described anecdotally in the journalistic literature ([Tiffany](#), 2021). Various pill variants emerged to signal the ideological splintering, here a brief overview:

- (4.1)** Blackpill schism (2011-2015): Under the moderation pressure and the increasing appropriation of the redpill terminology by various ideological actors, a splintering of the term into various morphological variations has been observed. Especially the term blackpill has become more common in mansphere and incel

communitas. This schism was motivated inasmuch by the broadening of the term in a semantic and ideological sense as it was driven by a need for internal differentiation. Now various variations exist:

- (4.2) Blue: The "default" state, thus comfortable ignorance of the redpill's claimed truths. Represents uncritical acceptance of feminist, progressive, or mainstream liberal values. Also, not yet awakened. Often interchangeably used with sheeple (sheep + people = controlled by elites or the system) or beta ($\neq$ alpha), cuck (a sexual practice in which someone's partner has sex with others; also, within BDSM and humiliation play). In a broader political usage, aligned with Democratic Party affiliation and institutional (over-)trust.
- (4.3) Black: fatalistic reinterpretation and extension of the redpill imaginary. Emerged around commenter Paragon on the Darlock blog; later popularized by the blog Omega Virgin Revolt. Associated with biological determinism and nihilism, often resulting in the glorification of violence and Darwinist ideals. No agentic dimension, often reflective of the impact on social media pressures on young men and the contemporary suicide crisis.
- (4.4) White: recovery from blackpill nihilism, thus optimism. Associated with podcaster Michael Malice and far-right accelerationist communities. Conveys hope for a coming upheaval that will restore the patriarchal or white nationalist order
- (4.5) Pink: female adaptation of the pill template with no fixed meaning. Often referring to women awakening to their oppression either within radical feminism, or conservative femininity (ex. Tradwife).
- (5) 'Being redpilled' has also globalized and spread into other languages, notably Germany, Brazil, Japan, and other European countries ([Noll for DW, 2025](#)). The word redpilled or related terminologies have appeared in various terrorist manifestos and/ or their online activity. In the *stricto sensu* these uses vary highly yet often yoking dissatisfaction with current social life and racial paranoia (i.e The Great Replacement). They mostly occur alongside the term's ideological broadening, allowing to link male loneliness to far right anxieties. Attacks with a clear reference include:
 - Halle, Germany 2019: Stephan Balliet kills two and tries to enter Synagogue. Balliet's manifesto and online presence explicitly used redpill terminology

alongside antisemitic and white nationalist content (Post ideological expansion)

- Toronto, Canada, 2018: Self-declared incel Alek Minassian drove a van into pedestrians, killing 10 and injuring 16. His Facebook post before the attack explicitly referenced Elliot Rodger and the "incel rebellion." Researchers document that Minassian used *redpill*-adjacent manosphere vocabulary, though his post uses incel rather than *redpill* directly (O'Donnell & Shor, 2022)
- Plymouth, United Kingdom, 2021: Jake Davison kills five and injures two. Davison explicitly described his mental state as "blackpilled" in his own words.

These figures also are represented within the dataset; either talked about or also as Redditors (CHECK). Other notable attacks (see below)

- (6) Since the term never provided a clear definition, even within the ingroup, an additional etymological form emerged. Including speculative definitions, counter discourse, and academic definition. Also gendered Inversions, and a retroactive reappropriation by the trans community, after the coming out of the Warwazik siblings. These definitions have become even more present after the quarantine and closing of various far right subreddits.

These entomological entries were compiled with the data set at hand and filtered Google Searches. It has also been retroactively confirmed by the literature and qualitative studies.

Sources

([Van Valkenburgh](#), 2018); ([Evans](#), 2018); ([Tiffany](#), 2021); ([Noll for DW](#), 2025); ([Hagen, Tuters & Wilson](#) for OILAB, 2020); ([O'Donnell & Shor](#), 2022);

After having established the term's lexicographic profile, the following part will trace the cultural history of *redpilled*, filling important gaps that cannot be traced empirically but only qualitatively via retracing the aforementioned semantic changes through individual comments from the dataset. The cultural history grounds the lexical structure in lived experiences and public controversies, providing important caveats to the term's semantic life. Previous studies have alluded to the mainstreaming of the term as an interplay of quantitative proliferation and constant renegotiation.

Discussion¹⁹

Lexicographic History on Reddit

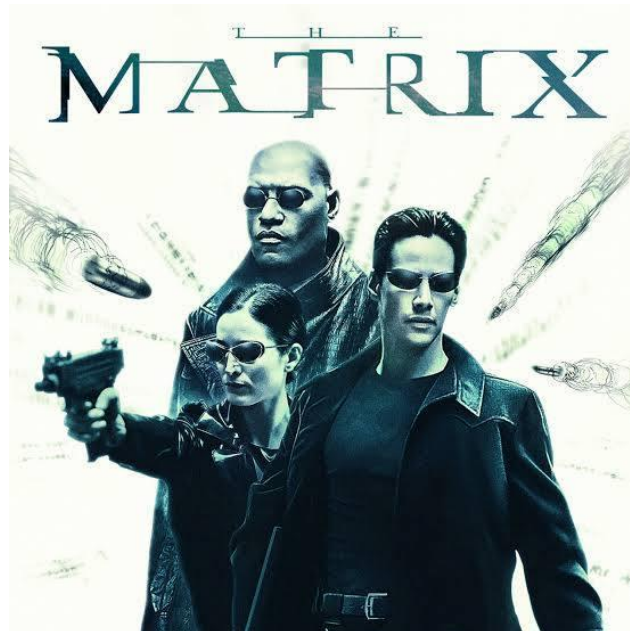


Photo 5 Taken from [TMDB \(The Movie Data Base\)](#)

To write a cultural history of the term redpilled starts with its cinematic inception. The now cult movie *The Matrix* (1999), reminiscent of late 90s cyber culture, serves as the most important cultural artifact *prima facie*. Directed by Lana and Lilly Wachowski, the movie follows Thomas Anderson, turned hacker Neo, who discovers that the world he inhabits is a simulated reality constructed by machines to keep humanity docile while harvesting their bodies as an energy source. Offered a choice between a blue pill that would return him to comfortable ignorance and a red pill that would reveal the truth, Neo chooses the latter, awakening into the devastated real world where a small band of rebels fights to free humanity from the machine's control. The film references a rich philosophical tradition spanning Plato's allegory of the cave, Descartes' evil demon, and Buddhist notions of illusion, framing the red pill as a symbol of painful but necessary epistemological liberation. It also toys with the tension between technology and human phenomenology, found in the philosophy of Martin Heidegger, Bernard Stiegler, or concurrently the accelerationist thinker Nick Land. Within the film's original context, the red pill is used as a purely cinematic device for dramatizing the choice between truth and comfortable ignorance.

¹⁹ Due to the word limitation less comments are featured than could have been



Photo 6 Taken from (Giorgio, 2025)

Ergo, the first use cases of the word were metaphoric experiments, often alluding to an extreme unescapable structuralism, mirrored by earlier theorists such as Louis Althusser. Yet cultural critics were also quick to point out that in a literary sense, the Matrix's red pill represents the conversion or entryway into a new ontological realm. A form of enlightenment or rite of passage through which an individual recognizes the oppressive force of the nowness and overcomes them through said recognition. Long before users would appropriate and later actually reify the pill's metaphor, this data set, in unison with OILAB's preliminary 2020 study, finds that the terms first mention remains metaphoric, yet ideologically neutral. The Redditor *u/acrophobia*. (2006, October 22) responds to a discussion about US-American English in an ironically reminiscent tone to French Theory:

"Eventually, everyone has to make peace with the depressing truth that the meaning of words is contextual and evolves according to its usage. I advise you pop the redpill sooner rather than later - it saves a lot of wasted energy."

The central cultural paradox of the term also lies in its creation. The Wachowskis, both trans women, though closeted at the time of filming, created *The Matrix* and its sequels as how Lily Wachowski put it: "(as) the desire for transformation, but coming from a closeted point of view" (BBC, 2020; citation from interview published on a website). Within film critic and the maker's own interpretation the red pill is sometimes seen as an allegory for hormone therapy and gender transition (Gender-confirming care: estrogen pills often being red). This tension might have informed the persistent gender play and association with sexual identity attributed to the early life of the term. *u/GnuYOW* comments in 2020, towards the end of the time frame where most comments become retrospective metacommentary about the impact and potential meaning of the term:

“I bet it drives the Wachowski “sisters” crazy that their movie became the metaphor for red pill philosophy and MGTOW.”

Only later would *redpill* incubate a broader meaning, in which it “*might lead to an epiphany about the rightness of white nationalism and/or the repudiation of feminism, multiculturalism, leftism, liberalism, and globalism, followed by the embrace of traditionalism, hierarchy, and inequality. There are thousands of red-pilling stories, each one slightly different, yet they all pivot around transformation*” (Taken from ([Hagen, Tuters & Wilson](#) for [OILAB](#), 2020), citing ([Stern](#), 2019, p.16). Towards the end of the data set in 2024, *u/bastrdsnbrokeknthings* even notes in *r/PurplePillDebate* “*I’m red-pilled as far as the sexual dynamics definition, but I’m definitely not aligned with the political right wing. As u/TheHumanDamage said, the definition of red pill has expanded in the past several years. (Now I’m wondering if the Wachowski (sisters?) intentionally chose political colors?)*” But what expansions of the term exist and when and where did they come to be?



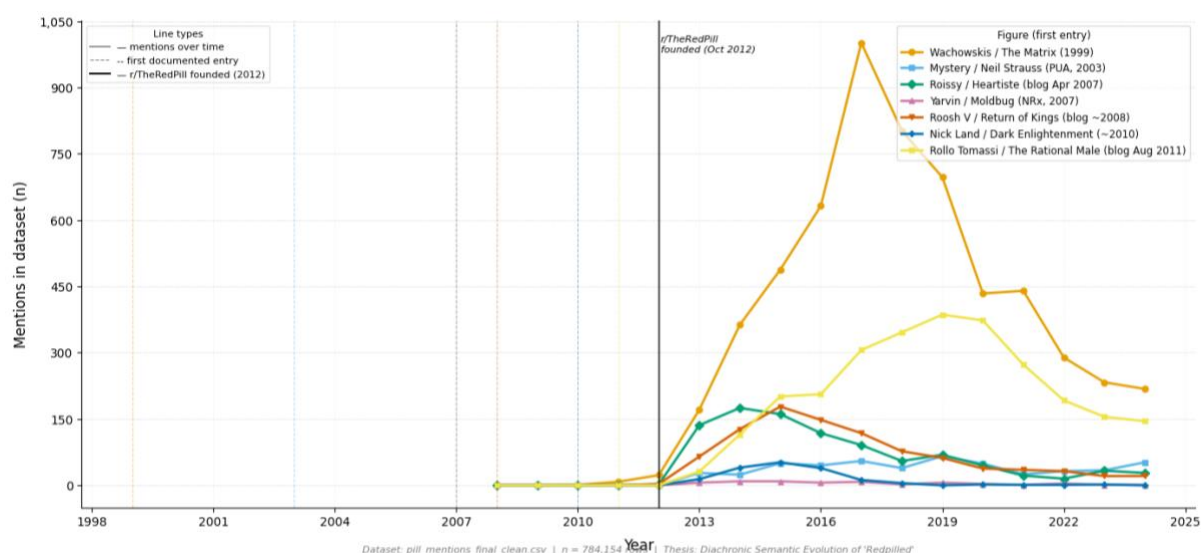
Photo 7 ([Waite](#), 2023)
Matrix-original use

Outside of the Reddit-based data set, the term *redpilled* appeared in various political writings (see table above). The period from 1999-2012 remains based on few unconnected sources that would later be foundational texts to both the manosphere and the political use case. Among many, within Curtis Yarvin’s deployment, who was then writing under the pseudonym Mencius Moldbug, being *redpilled* “*roughly meant gaining independence from the standard liberal narratives*” ([Shullenberger](#), 2021). 56 comments from the data set mention him directly in connection to *redpill*, especially in its early days. Others include the early blogosphere (Return of Kings, The Rational Male, Roosh V's blog), and other redditors. All external sources are used retrospectively, which might imply a desire to retrace the word’s origin or as a projection of their political project onto its contemporary meaning. The user *Jon Frost* describes Curtis Yarvin as a necessary entry point to the *redpill* culture as

following:

“(…) I was a Libertarian/Objectivist in my early teens, so I got used to having unpopular opinions and thinking everyone around me was nuts from a young age. Intellectually, reading Moldbug was a major step for me. As far as women, dating a few girls with boyfriends in my sophomore year of undergrad was an eye-opener on the darker side of female nature. 2) Hah. Fortunately I haven't lost any friends. Drifted away from a few, sure. As for what the future holds, I know that a lot of my good friends' gfs and fiancées are not my #1 fans. This will inevitably lead to distance as we age (…)”

Various pre-2012 redpill precursors, those who have engaged with or used the term prior to its *manosphere* definition and are also recognized by the literature (see below), are featured throughout the dataset. The mention frequency remains low however, and declines towards the end. The Wachowski siblings are, after President Donald J. Trump and various Reddit native users, the most common private individuals semantically linked to *redpilled*.



Graph from Dataset 12 Pre-2012 Redpill Precursors with Mention Frequency Over Time

Manosphere use

While the mansphere is most often associated with *redpilled* by the literature, it did not originate the term. Equally, a political orientation of the apothegm is often premeditatively imputed to the Incel-movement. Multiple factors coalesced to create the plurosemantic terminology we know today. However, even prior to the construction of this dataset, the notorious quote attributed to u/pk_atheist, pseudonym of Robert Fisher, a New Hampshire resident turned men's rights activist and later state Representative in the New Hampshire House, informs the popular cultural understanding of the term's early ideological valence.

“Welcome to the Red Pill I'm going to discuss briefly what my intention is for this subreddit. I'm Desmond, and I've been active in both the Men's Rights and the Seduction subreddits. They're both wildly popular subs, but both have major failings that I've slowly identified. They

both operate subtly under the feminist imperative. Group-think at both tend to fail to grok the importance of coming to terms with objective reality - something the manosphere has termed "taking the red pill." I haven't seen serious discussion on sexual strategy that goes beyond the pickup. We've heard the common trope in many men's subreddits: "Not all women are like that" That's true. Don't get me wrong. But something it undermines is honest discussion about evolutionary psychology, and society in terms of what a sexual life is like for a man. "How is that?" I hear you ask. It's simple. Let's describe a problem. (..)"

Many early comments align with this foundational definition, of men realizing in both a sincere and ironic way that feminism is harming men, though they would soon be expanded by notions of the sexual market and a hardening, increasingly aggressive PUA strategy register, alongside a parallel current of communal bonding around isolation and self-described masochisticism, as in *u/cuckoo99's* 2013 *r/TheRedPill* comment that *"Askwomen is a joke. I'm a masochist and cuckold fetishist and I've posted there before, and they absolutely hated me, not because of my fetish, but because I posted 'red-pill' insights"*. *r/TheRedPill* was, however, no hermetic invention. Its early user base overlapped substantially with the existing MRA and PUA communities, and *r/MGTOW* operated practically as a subset of *r/MensRights* through 2012-2013 (Ribeiro et al., 2021). Ribeiro et al. note that since its creation in October 2012, *r/TheRedPill* and its adjacent subreddits grew at a fast pace, becoming the most active community by the end of 2013 in terms of number of posts, and the most popular community by the middle of 2014 in terms of number of active accounts. The subreddit thus consolidated and rebranded existing manosphere discourses under a new metaphor rather than creating them ex nihilo, a process visible in the dataset through *u/pk_atheist's* 2013 openly recruitment-oriented presence in adjacent subreddits, where he writes in *r/MensRights*:

"This might be a little too red pill for this subreddit. Although it really matches what I understand about female mating schedules.. you might be better received over at /r/theredpill. We talk about mating and sexual strategy in the light of feminism."

This cross-subreddit traffic was not uncontested. There is also evidence of strong in-group policing in the older men's-rights spaces, where users sought to contain ideological cohesion, often using *redpill* to exclusively refer to the *r/TheRedPill* exclusive semantic use. It aims at retaining a productive discursive space for pro-male advocacy and to expel its more virulent strain. As *u/CloverFuchs* writes in *r/MensRights* in 2013:

"Kindly get the fuck out. You are to the MRA as SJW and Feminazis are to Feminism. We don't need your kind here and we certainly don't want you. Try r/redpill if you'd like to circlejerk about how much you hate feminists and the bigoted, derogatory things you'd like to say about their appearance. I, personally, would like our movement to be taken seriously, and it's fuckheads like you ruining it for the rest of us."

The internal fragmentation visible throughout the term's early use can also be observed in the emergence of *r/RedPillWomen*. A companion subreddit active and growing by late 2013, *r/RedPillWomen* (RPW) brought together women who accepted the *redpill*'s gendered premises and sought to enact them, embracing traditional femininity, submission to male partners, and the rejection of feminist principles. As *u/TempestTcup* writes in 2013: "*Maybe a good red pill strategy for a woman, other than just looks, is to figure out how to add value to her relationship. Keep the house nice, keep her man well fed with nutritious food, keep things interesting in bed, etc.*" These communities idealized a 1950s white, heteronormative American family model. Men could not post on RPW unless verified as 'Red Pill Men' through participation on the main TRP subreddit, creating a gatekeeping mechanism that reinforced the ideology's internal hierarchy and a gendered radicalization framework. Yet this same mechanism gatekept new semantic incubatory spaces, fixing the meaning of *redpilled* within a tightly controlled discursive perimeter. Anecdotally from the comments themselves, *redpilled* at this stage still primarily denotes an acknowledgment and realization of the ideology proposed by *r/TheRedPill*, though it also signals a clear opposition to progressivist social forms more broadly, exceeding even *pk_atheist*'s original framing, which placed responsibility on men for failing to become "real men." By 2013–2014, then, *redpilled* had come to mean, above all, awakened to the truth that feminism has rigged society against men. *u/Brandwein* expresses his frustration in late 2014 on *r/MGTOW*:

"Yep. RedPill was once the dominant term for all of the manosphere. Now it is taken by a cult of elitist PUAs that just abuse the TRP knowledge for social and sexual power. Talking crap about self-improvement, but most of it is just about building muscles, only seldom about fun, knowledge, freedom and other important things. They often reject all concepts of self-fulfillment that are not "alpha."

#Gamergate: Politicizing the Redpill

The term changed substantively after August 2014, because the online campaign #Gamergate functioned as a mass entry point through which tens of thousands of young men, many of them previously politically disengaged gamers, encountered the *redpill* framework for the first time and learned to apply it outside the sexual-strategic register established by *pk_atheist* and the early *r/TheRedPill*. Gamergate began with "The Zoe Post," a lengthy blog entry published on August 16, 2014 by Eron Gjoni, the ex-boyfriend of game developer Zoë Quinn, which falsely insinuated that Quinn had exchanged sexual favors for favorable coverage from a games journalist. The accusation was baseless, since the journalist in question, Nathan Grayson, had only briefly mentioned Quinn's work, and had done so before any relationship

with Quinn began ([Jason for Boston Magazine](#), 2015). Yet the post ignited a coordinated harassment campaign against Quinn and other prominent women in the gaming industry, conducted across 4chan, 8chan, Reddit, and IRC. Actor Adam Baldwin coined the hashtag *#GamerGate* on August 27, 2014, and Milo Yiannopoulos amplified it through Breitbart News, becoming one of its most visible faces. Within the dataset, this moment is most visible through r/KotakuInAction, the principal Reddit hub for *Gamergate*-aligned discourse.

However, the term changed only gradually because Gamergate and the manosphere *redpill* were not initially aligned. In their earliest stages the two communities perceived each other as distinct, and Gamergaters often resisted being conflated with r/TheRedPill proper. As u/Aurondarklord writes on r/KotakuInAction in 2016: *"Dude, you're not even a GamerGater, you're some redpill who wants to come to our sub and demand the right to co-opt our mission"*. Yet, the real boundary policing did not hold, because the two discourses shared a structural template, an awakening to a hidden corruption sustained by a progressive elite, that made cross-pollination almost inevitable once the harassment campaigns drew the same demographic into both spaces. The dataset registers this slippage in 2015, when u/despoticVeracity writes on r/PurplePillDebate that

*"'Red Pill' is not a term exclusively used by /r/TheRedPill. People on fringe boards, such as the *chans' /pol/, also use the term. It signifies an awakening from a commonly held belief in society. You can be 'redpill' on conspiracies, on Jewish influence, on race-relations or on gender interactions. /r/TheRedPill exclusively deals in the latter. (...) TRP does not espouse any views on conspiracy theories, Jews or race-relations (aside from the impact on gender and sexuality). Its only advice would be 'don't say that shit in public, 'cause it'll get you ostracised"*

The comment is doubly significant, because it documents both the existing semantic spread of redpill beyond the manosphere by 2015 and r/TheRedPill's active attempt to disavow that spread, suggesting its moderators recognized the reputational and platform risks of being associated with /pol/ and Gamergate's racially charged discourse. Pk_atheist, at that point a NH state politician, was equally trying to cover up his secret manosphere *alter ego*.

Secondly, the term changed because Gamergate served as a proving ground for digital tactics that would later be redeployed by the alt-right in the 2016 election cycle: coordinated harassment, doxing, meme warfare, and narrative flooding. Steve Bannon, then heading Breitbart, explicitly recognized the political potential of the disaffected young male demographic Gamergate had mobilized, later remarking that through Yiannopoulos he could activate an army²⁰ of gamers and internet trolls. Banon for example treated Gamergate as a

²⁰ Cf. Also [USA TODAY Tech](#).

template for online political organization that could be scaled into the MAGA movement. Once those tactics had a name and an audience, the *redpill* metaphor traveled with them, no longer tied to the sexual marketplace but adaptable to any awakening narrative the same audience could be persuaded to undertake²¹.

The term changed, lastly, because Gamergate bridged the manosphere's gender-focused *redpill* with the emerging alt-right's political *redpill*, fusing anti-feminism with anti-progressivism. The composite meaning that would come to dominate the Trump era, awakened to the supposed failures of liberalism, feminism, diversity, and political correctness simultaneously, was assembled here. The dataset registers this composite emerging in real time. From 2015 onward, references to Jewish media influence, "sheeple," "goy cattle," and similar tropes begin to surface in adjacent spaces, despite r/*TheRedPill* and the classical manosphere attempting to distance themselves from such content, as the u/*despoticVeracity* comment cited above demonstrates. By 2016, that disavowal had proven unsustainable. Openly political articulations of the redpill had begun to appear within manosphere spaces themselves. As u/*The_DogeWhisperer* writes on r/*asktrp* in 2016:

"What does The Red Pill think of the current migrant situation in Europe? Millions of Muslim refugees have immigrated into Europe for one reason or another. We know that the current numbers of migrants in countries like Britain commit crimes [at the same rate as the entire population] despite comprising [less than 4.5%] of the population. We also know that descendants of non-Western immigrants are [218% more likely to commit crime]"

The framing of the question is itself the evidence, since by 2016 asking r/*asktrp*'s view on European migration is treated as a coherent question, presupposing that the *redpill* framework is now expected to have political and racial-demographic positions and not only sexual-strategic ones. The boundary u/*despoticVeracity* had drawn the year before has effectively dissolved.

Donald Trump & Redpilling

Donald Trump's 2016 presidential campaign and subsequent victory constituted the largest single inflection point in the mainstreaming of redpilled, because three distinct forces converged within roughly eighteen months to shift both the term's referent and its grammatical behavior. The shift is registered in the dataset both quantitatively, through the explosive growth of r/*The_Donald* (173,957 comments by the end of the data window, the second largest sub in the corpus), and qualitatively, through a measurable change in how users deploy the term.

²¹ Cf. Meme Wars



Photo 8 Taken from (Tiffany, 2021)

The term changed, first, because Steve Bannon's Breitbart News served as the institutional bridge between the discursive infrastructure built during Gamergate and mainstream conservative politics. Bannon, who described Breitbart at the 2016 Republican National Convention as "the platform for the alt-right," became Trump's campaign CEO in August 2016 and was subsequently appointed White House Chief Strategist. Breitbart's editorial strategy drew heavily on *redpill* rhetoric, framing progressive consensus as a manufactured illusion that Trump's movement was now shattering. The site functioned as a conduit between 4chan culture, manosphere ideology, and mainstream conservative politics, channeling the digital tactics and semantic structures forged during Gamergate into electoral politics.

The term changed, second, because the Trump era produced a critical grammatical shift in how *redpill* was used. Where *pk_atheist's* 2012 founding post had described "*taking the red pill*" as a passive, individual act of awakening, by 2016 *redpill* was increasingly being used as a verb, and a transitive one at that. One no longer merely took the redpill but could *redpill* another person, actively converting them to a set of right-wing beliefs. This shift from intransitive to transitive use marks the term's evolution from describing personal enlightenment to describing proselytization. Trump supporters, and later QAnon adherents, spoke of "redpilling normies," initiating outsiders into their shared belief system. The dataset captures this shift in real time. As *u/jesseale* writes on *r/The_Donald* in November 2016 in an already seen comment:

"Trump support is surging. Expect 60% by Tuesday. GO VOTE! WE NEED A RECORD BREAKING LANDSLIDE TO OVERCOME THIS RIGGED ELECTION. START RED PILLING YOUR FRIENDS AND FAMILY ON THE HORRORS OF HILLARY CLINTON"

The imperative grammar is the evidence: *redpill* is no longer something one undergoes but something one *does to others*, and at scale, with electoral stakes attached. The metaphor has been operationalized as a campaign tactic.

The same Trump moment also produced an awareness, within manosphere spaces themselves, that the term's mainstreaming had political consequences for the communities that had originated it. As *u/PabloEscoba* writes on *r/MGTOW* in 2016:

"Prepare for a MGTOW and redpill crackdown gentlemen. Now that the powers that be have witnessed first hand how alternative media aided Trump in winning the presidency we should prepare for a MGTOW and redpill assault on all fronts. Look at the exponential rise of mainstream shaming articles about MGTOW. Look at the outright ban of the redpill documentary in Australia. Censoring us is no longer a viable strategy. They will no longer allow us to just be. We are too dangerous to ignore"

This comment documents the manosphere's own self-perception of having entered mainstream political relevance through Trump's victory. Second, it anticipates the platform-moderation arc that would in fact unfold over the following five years, with *r/Incels* banned in 2017, *r/TheRedPill* quarantined in 2018, *r/The_Donald* banned in 2020, and *r/MGTOW* itself banned in 2021. The mainstreaming of *redpill* and the deplatforming of its original communities turn out to be two sides of the same dynamic.

The term changed, also, because the QAnon movement, which emerged in late 2017, adopted *redpill* language wholesale and stripped it of its remaining gendered content. In QAnon's framework, *redpilling* meant awakening to an elaborate conspiracy theory about a secret cabal of elites engaged in child trafficking, with Donald Trump cast as the hidden hero waging war against them. The *redpill* metaphor allowed QAnon to frame its conspiracy theories as hidden truths that the mainstream media was suppressing, the same epistemological structure that Yarvin had pioneered a decade earlier, now applied to content vastly more extreme and demonstrably false. The dataset registers this shift through *r/greatawakening*, the principal QAnon hub on Reddit until its ban in September 2018, and through the broader substrate of *r/conspiracy*, where the *redpill* metaphor predated and outlasted Q. The Q-era *redpill* is consistently framed as effortful labor rather than passive epiphany, with users casting research, sharing, and recruitment as the moral substance of the awakening itself. As *u/bgny* writes on *r/conspiracy* in March 2020:

"How is doing all the work of researching, sharing, redpilling, and motivating passive? Was the vast amount of research done exposing the cabal over the last few years a passive activity? The Great Awakening is a positive and ACTIVE truth movement. The blackpilled are the ones doing nothing but complaining, since they are mired in nihilistic hopelessness. (...) We have the unsealing of the federal indictments of Epstein, Raniere, Mack, Smolette, thousands of pedophiles and child traffickers"

The comment performs the same transitive grammar visible in the Trump-era political *redpill* yet explicitly contrasting to the *blackpilled* community along an axis of agency. It is also at

this stage where the Epstein conspiracy is more commonly and nonchalantly evoked, linking Trump's first presidency to a revelatory time in uncovering the hidden truth of this digitally native tale. The word *redpill* moves closer to its original meaning of a matrixed world.

The QAnon *redpill* also retained, however, an open channel to the racially explicit discourse that *r/TheRedPill* had attempted to disavow in 2015. As *u/winterfjell* writes on *r/altright* in 2017: "*I think Goy²² is fine here, it's used ironically. It's a hidden red pill for the curious and a term of membership / belonging for us.*" The comment is significant because it frames the *redpill* not as a single act of awakening but as a layered system of "hidden" pills delivered through ironic in-group language, with antisemitic shibboleths functioning simultaneously as group-recognition markers and as recruitment devices for "the curious." The same logic, scaled to longer narrative form, appears in *u/ScythianRabbit's* 2019 *r/DebateAltRight* post, which uses a video-game playthrough of *Stellaris* as an extended genocidal allegory for ethnonationalist *redpilling* against "*Jewry (that) undermin(s) their globalist dreams (by) (m)isreading the internet and pushing the goy too far*". The user contemplates a "*space fascist human*" empire that conquers a "*diverse cesspit*" planet and deports its non-human inhabitants to a peripheral "*shithole*" world.

By the close of the Trump era, then, *redpilled* had become a mainstream political adjective and a transitive verb of proselytization. The term now signified MAGA-conservative or right-wing political awakening broadly, detached from any single community or specific doctrine, and could be applied to anything from electoral politics to media criticism to conspiratorial cosmology. The semantic perimeter that *pk_atheist* had drawn in 2012 and that *u/CloverFuchs* and *u/despoticVeracity* had attempted to defend in 2013 and 2015 respectively had finally and definitively dissolved.

²² = non Jew

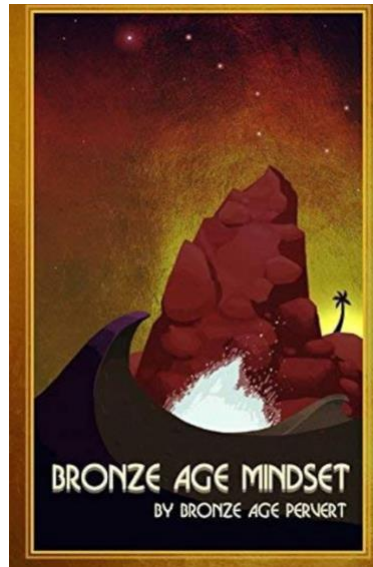


Photo 9. Cover of Bronze Age Mindset by Bronze Age Pervert (Costin Alamariu, 2018), a self-published book widely circulated in alt-right and manosphere reading lists and frequently cited as a foundational text within the broader incel and identitarian canons. References to the book recur throughout the dataset, where it tends to appear alongside thinkers such as Curtis Yarvin, Nick Land, and Guillaume Faye as part of a shared canon of redpill-adjacent reading. As u/Durango_Sequel writes on r/DebateAltRight in 2019:

"An easy intro to the basics of Identitarian thought is Archeofuturism by the recently late Guillaume Faye. If you want a meme-y and easy reading Nietzsche, you can pick up Bronze Age Mindset. (...) Austrian/an-cap economics can inform more authoritarian systems, right on up to National Socialism (...) The 'redpill' isn't only racial and such, but also economic".*

The comment is doubly significant, since it both confirms the book's position within the redpill canon and explicitly extends the metaphor into the economic register, gesturing toward Austrian-school and libertarian thought as further redpill entry points.

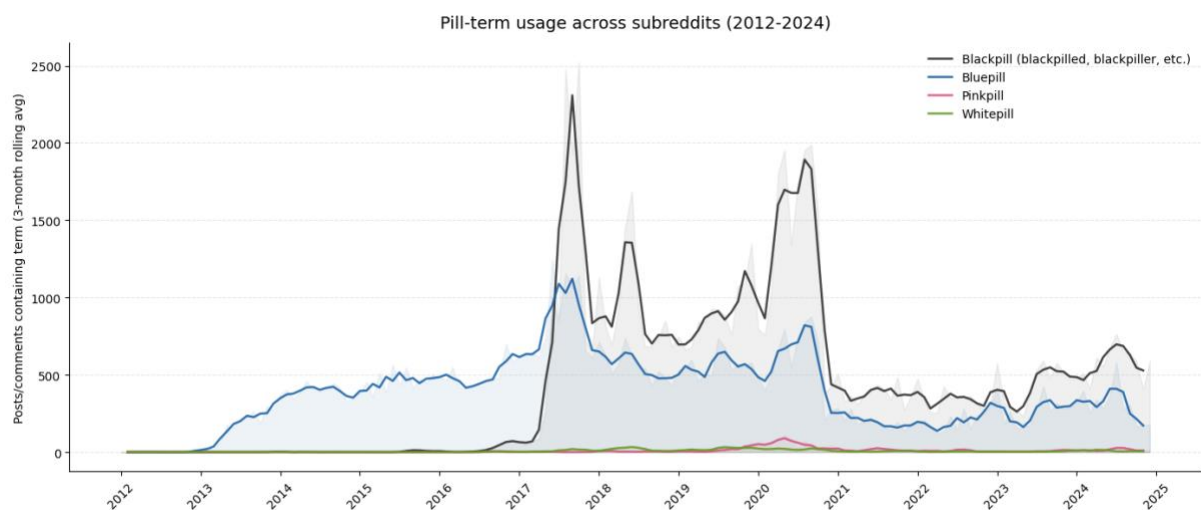
Manosphere Terrorism

A particular caveat for any cultural history of redpilled is the possibility of tracing the term's relationship to a series of mass-violence events that punctuate the period covered by the dataset. Between Elliot Rodger's 2014 Isla Vista attack and Lyndon McLeod's 2021 Denver shootings, at least six attacks have been linked to incel or *redpill*-adjacent radicalization. The dataset registers these events not only through retrospective discussion but, in some cases, through the participation of perpetrators themselves in the communities under study²³. Rodger is repeatedly invoked as a *redpill* reference point, including by users who explicitly frame him as catalytic to their own awakening. As u/geocitiesuser writes on r/MGTOW in 2016:

"Elliot Rodger was a prophet that helped Redpill me. His manifesto laid out exactly a lot of faults in society and called women out for what they are."

²³ Also emergence of expressions like „going er“ = going Elliot Rodger to condone mass violence in the name of redpill communities

At the same time, the dataset records the proliferation of an escalating pill typology, with blackpill, whitepill, and ironpill emerging as named alternatives that index different attitudes toward the agency the original redpill had implied. This typological multiplication accelerates after 2017, particularly within incel-adjacent spaces, and is captured in pithy, often retrospective formulations such as "*Red pill is wank. Black pill is truth,*" posted to r/Braincels in 2017. The *redpill*, in other words, becomes the baseline against which more extreme positions define themselves, and the dataset therefore captures both the term's mainstreaming and the simultaneous radicalization of the communities for which it had once been the most extreme designation available.



Graph from Dataset 13 Pill-term usage across all subreddits (2012-2024)

The *redpill* is also globally distributed throughout the dataset, with German, Brazilian, British, and other non-American users appearing as a steady minority of the corpus. As *u/Pretorius1836* writes on *r/DebateAltRight* in 2018, in response to a thread asking "How were you 'Redpilled'?":

"I am a German from ex East Germany (Brandenburg to be precise). My parents are all Prussian refugees who had their homeland and culture taken from them by the Allied powers, so I always was suspicious of 'the good guys' (...)"

The comment projects the *redpill* awakening narrative onto a specifically German revisionist memory politics. Similar adoption is visible in German (*redpilled sein*) and Brazilian Portuguese (*redpillado*, *tomar a pílula vermelha*), where the term is typically loaned rather than translated. While these forms are not present in the entirely English dataset, speakers and users of these translations are. Said non-translation is itself analytically significant, since it suggests that participation in the global *redpill* discourse is partly performed through the choice to retain the English form.

Counter-Discourse, Virality²⁴, and Open Signifier

The shift in the term's referent over the 2018–2024 period cannot be separated from the platform-moderation arc that hollowed out the communities that had originally stabilized its meaning. The bans of *r/altright* (2017), *r/Incels* (2017), *r/greatawakening* (2018), the quarantine of *r/TheRedPill* (2018), the ban of *r/The_Donald* (2020), and the ban of *r/MGTOW* (2021) successively removed the institutional homes within which each of redpill's semantic registers had been incubated. When the communities that had defined a meaning disappeared, that meaning lost its anchor, and the term drifted into the ambient discourse of surviving spaces, where it became increasingly available for ironic, diagnostic, and counter-discursive reuse. The dataset registers this in two simultaneous trends. First, the surviving subreddits are disproportionately counter-discursive (*r/exredpill*, *r/TheBluePill*, *r/IncelExit*, *r/niceguys*), debate-structured (*r/PurplePillDebate*), or critically observational (*r/QAnonCasualties-adjacent threads on r/conspiracy*). Furthermore, sincere uses of the term within these spaces are increasingly framed as harmful to the user's own life, as in *u/Standard_Version610*'s 2023 comment on *r/IncelExit*:

"I CAN just ignore all the Relationship Advice and Gender War bullshit! I think my brain is filled with too much relationship advice. Too much power play, redpill, blackpill philosophy crap, that it's actually actively harming how I interact with the women I like"

Content moderation (H2), while often analyzed in terms of speech and access, also functions as a form of *semantic governance*. The disappearance of the communities that had stabilized particular meanings of redpill produced a term that survives lexically but has lost the discursive scaffolding that gave it ideological precision. By 2024, redpill is more ambiguous than at any earlier point in the data set, precisely because the inventors of its semantic use cases are no longer present to police it. This allows for an increased plausible deniability, which is already inherent to the open signifier. The term floats freely, available to be redeployed by mainstream conservative commentators, by ironic posters, by deradicalization spaces, and by retrospective metacommentary.

Across the period studied, redpill has functioned not as a single concept but as a structural template, an awakening narrative whose content has been progressively reclaimed by successive communities. The categories identified in this analysis (manosphere, MAGA, QAnon, incel, contested debate, counter-discourse) and the registers that cross-cut them

²⁴ After 2018 the term was picked up by major celebrities such as Elon Musk, Kanye West, Andrew Tate or Candace Owens with *Redpill Black* to show their conversions to conservative viewpoints. However, these are less often featured in the data set and are viewed as outside occurrences.

(neutral, ironic, ideological, political, conspiratorial, extremist, transitive) are not stages of a single evolution but a stratigraphy of overlapping uses, each laid down in a particular community at a particular moment, each persisting in some form even after that community has been moderated or has migrated elsewhere. The last table maps these meanings onto the subreddits that incubated them, and shows that no single user, sub, or year owns the term's definition.

Manosphere redpill	Redpilled				
	MAGA (Pro-Trump)	QAnon	Incel	Contested debate	Counter-discourse
Awakening to the "truth" that feminism oppresses men; largest use case with contested debate; both ironic and sincere uses. Becomes more sincere towards end, hence moderation	Awakening to anti-establishment Trump-aligned politics. May be gendered when Trump is framed as anti-feminist or pro-Men's Rights	Degendered awakening to hidden elite cabal. Emerges from MAGA discourse. Focuses largely on Q (sometimes =Trump)	Fatalist branch: awakening reveals nothing can be done; often linked to Incel terrorism	Meta-frame where the metaphor itself is contested; only semantic use that references <i>The Matrix</i> (1999) Regularly; openly speculates about semantic meaning for sincere users	Diagnostic, parodic, or exit-oriented use
<i>r/TheRedPill</i> <i>r/asktrp</i> <i>r/MGTOW</i> <i>r/RedPillWomen</i> <i>r/marriedredpill</i>	<i>r/The_Donald</i> <i>r/4chan4trump</i> <i>r/milliondollar</i> <i>extreme</i> <i>r/DebateAltRight</i> <i>ht</i> <i>r/altright</i>	<i>r/conspiracy</i> <i>r/greatawakening</i>	<i>r/Incels</i> <i>r/IncelsWithoutHate</i> <i>r/Braincels</i>	<i>r/PurplePillDebate</i> <i>r/MensRights</i> <i>r/KotakuInAction</i> <i>r/kotakuinaction2</i>	<i>r/TheBluePill</i> <i>r/exredpill</i> <i>r/niceguys</i> <i>r/IncelTears</i> <i>r/againstmensrights</i> <i>r/IncelExit</i>
Always present	From 2015-2016 onwards, mostly halts after <i>r/The_Donald</i> is quarantined and banned	2016-	2018-onwards	Always present, yet especially after 2019, when all radical places of semantic origin were moderated	After 2014

Redpill is, in this sense, a word that has outlived each of its makers. It began as a cinematic image of a single, brave choice between truth and illusion, was taken up as a doctrine by men who were sure of what the truth was, was politicized into a campaign tactic, conspiratorialized into a parallel cosmology, weaponized in acts of mass violence, mocked in memes, translated into German and Portuguese, and finally diagnosed by the very people it had once promised to save. What remains, at the end of the data set, is the original Wachowski gesture freed of any one community's claim on it. A hand offering a pill, an open question about which truth one is awakening to, and a quiet recognition that the choice, once made, was never as singular as the metaphor pretended.

Recommendations²⁵

Domain	Lesson	Recommendation
Content moderation and semantic governance	Bans and quarantines (<i>r/Incels</i> 2017, <i>r/TheRedPill</i> 2018, <i>r/greatawakening</i> 2018, <i>r/The_Donald</i> 2020, <i>r/MGTOW</i> 2021) did not eliminate <i>redpilled</i> . They dissolved the communities that anchored it. The term survived lexically but lost the scaffolding that gave it ideological precision, becoming more available to plausible deniability and external reappropriation. Moderation is never neutral excision; it is a productive intervention whose consequences outlast the communities removed.	Treat moderation as semantic governance, not speech regulation. Pair bans with longitudinal lexical monitoring. Support counter-discursive communities as legitimate post-moderation infrastructures of meaning-making. Time interventions to the S-curve: early incubation is more tractable than peak diffusion, when removal accelerates drift rather than containing it.
The aesthetic of democratic language	Digitally native neologisms operate under the Written Speech Paradox. This structurally favors irony and plurosemy. <i>Redpilled</i> functioned sincerely, ironically and neutrally. The gap between <i>adequate</i> and <i>correct</i> use became the structural space of plausible deniability.	Build a literacy of constraint-shaped registers in education, journalism, and democratic institutions. Equip citizens to ask not what a term means in the abstract but in which community, at which phase of

²⁵ Due to the word limitation presented within a table format

	The Habermasian deliberative ideal assumes a referential stability that constraint-shaped registers erode.	its drift, and in which register it is operating. Resist policing digital language back into analogue norms without treating irony as innocent.
The future of the internet	The Internet's text-centered architecture, combined with anonymity and algorithmic curation, produces semantic acceleration without semantic stabilization. The decoupling of ideological potency from semantic charge is structural, not accidental, and will intensify as AI-generated content enters the same semiosphere.	Design platforms and research infrastructures around perpetual semantic flux. Invest in open archival infrastructures; the Pushshift case shows what is lost when archives are gated. Fund diachronic linguistic research as public-interest work. Develop affordances that make register and provenance legible to users without imposing a single authoritative reading.

Moritz Schwenger

25285 words, excluding the dictionary table

References

Sources

Made using Zotero bib APA 7TH EDITION Citation:
<https://zbib.org/7e26fe6b3ab54138b4b288b8473ed912>

All sources are also hyperlinked throughout the text. For further questions or potential error due to the automated process via Zotero bib, please contact moritz.schwenger@sciencespo.fr

Åkerlund, M. (2022). Dog whistling far-right code words: The case of ‘culture enricher’ on the Swedish web. *Information, Communication & Society*, 25(12), 1808–1825.

<https://doi.org/10.1080/1369118X.2021.1889639>

Alamariu, C. or Bronze Age Pervert (2018). *Bronze age mindset - pervert, bronze age:*

9781983090448 - abebooks. <https://www.abebooks.com/9781983090448/Bronze-Age-Mindset-Pervert-1983090441/plp>

Aleksic, A. (2025). *Algospeak: How social media is transforming the future of language*. Random House LLC US. (Original work published 2025) ISBN 0593804074

Androutsopoulos, J. (2011). Language change and digital media: A review of conceptions and evidence. In *Standard languages and language standards in a changing Europe* (pp. 145–159). Novus Press. <https://jannisandroutsopoulos.net/wp-content/uploads/2011/11/language-change-and-digital-media-preprint.pdf>

Baele, S., Brace, L., & Ging, D. (2024). A diachronic cross-platforms analysis of violent extremist language in the incel online ecosystem. *Terrorism and Political Violence*, 36(3), 382–405.

<https://doi.org/10.1080/09546553.2022.2161373>

Bankov, K. (2022). The digital semiosphere. In K. Bankov (Ed.), *The Digital Mind: Semiotic Explorations in Digital Culture* (pp. 3–23). Springer International Publishing.

https://doi.org/10.1007/978-3-030-92555-0_1

Baudrillard, J. (2007). *Simulacres et simulation* (Reprint from 1981). Éditions Galilée. ISBN: 2718602104

BBC, n. a. (2020, August 7). The Matrix is a ‘trans metaphor’, Lilly Wachowski says. *BBC*.

<https://www.bbc.com/news/newsbeat-53692435>

- Berthon, P., Pitt, L., & Watson, R. T. (2000). Postmodernism and the web. *Technological Forecasting and Social Change*, 65(3), 265–279. [https://doi.org/10.1016/S0040-1625\(99\)00067-0](https://doi.org/10.1016/S0040-1625(99)00067-0)
- Blank, A. (1999). Why do new meanings occur? A cognitive typology of the motivations for lexical semantic change. In A. Blank & P. Koch (Eds), *Historical Semantics and Cognition* (pp. 61–90). DE GRUYTER. <https://doi.org/10.1515/9783110804195.61>
- Blank, A., & Koch, P. (2013). *Historical semantics and cognition*. De Gruyter Mouton. <https://doi.org/10.1515/9783110804195>
- Bloomfield, L. (1933). Linguistic aspects of science. *Philosophy of Science*, 2(4), 499–517. <https://doi.org/10.1086/286392>
- Bourdieu, P. (1991). *Language and symbolic power* (J. B. Thompson, Ed.). Harvard University Press. ISBN 9780674510418
- Bratich, J. Z. (2023). Redpilling and the archaic roots of patriarchal post-truth. In J. Harsin, *Re-thinking Mediations of Post-truth Politics and Trust* (1st edn, pp. 89–108). Routledge. <https://doi.org/10.4324/9781003388975-5>
- Bujdei-Tebeica, V. (2023). Neoliberalism and the red pill: How andrew tate used populist discourse to spread online misogyny. *Civil Szemle*, 4, 5–12. https://www.researchgate.net/publication/376488452_Neoliberalism_and_the_Red_Pill_How_Andrew_Tate_Used_Populist_Discourse_to_Spread_Online_Misogyny
- Campbell, L. (2013). *Historical linguistics: An introduction* (3. ed). MIT Press. ISBN 0748675590
- Cannizzaro, S. (2016). Internet memes as internet signs: A semiotic view of digital culture. *Sign Systems Studies*, 44(4), 562–586. <https://doi.org/10.12697/SSS.2016.44.4.05>
- Chandler, D. (2009). *Semiotics: The basics* (2. Edition, reprinted). Routledge. https://www.researchgate.net/publication/263768124_Chandler_Daniel_2007_Semiotics_The_Basics_2nd_edn_London_Routledge

- Copland, S. (2020). Reddit quarantined: Can changing platform affordances reduce hateful material online? *Internet Policy Review*, 9(4). <https://doi.org/10.14763/2020.4.1516>
- Crystal, D. (2001). *Language and the internet* (1st edn). Cambridge University Press.
<https://doi.org/10.1017/CBO9781139164771>
- Crystal, D. (2006). *Language and the internet* (2nd edn). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511487002>
- Del Tredici, M., & Fernández, R. (2018). The road to success: Assessing the fate of linguistic innovations in online communities. In E. M. Bender, L. Derczynski, & P. Isabelle (Eds), *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 1591–1603). Association for Computational Linguistics. <https://aclanthology.org/C18-1135/>
- Denison, D. (2003). Log(Ist)ic and simplistic S-curves. In R. Hickey (Ed.), *Motives for Language Change* (1st edn, pp. 54–70). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486937.005>
- Derrida, J. (2016). *Of grammatology*. Johns Hopkins University Press.
<https://doi.org/10.56021/9781421419954>
- Derrida, J., & Johnson, B. (1981). *Dissemination*. University of Chicago Press.
<https://doi.org/10.7208/chicago/9780226816340.001.0001>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- Eco, U. (1976). *A Theory of Semiotics*. JSTOR. <https://www.jstor.org/stable/j.ctt16xwcfj>
(Original work published Indiana University Press)
- Eger, S., & Mehler, A. (2016). On the linearity of semantic change: Investigating meaning variation via dynamic graph models. *Proceedings of the 54th Annual Meeting of the*

Association for Computational Linguistics (Volume 2: Short Papers), 52–58.

<https://doi.org/10.18653/v1/P16-2009>

Egger, R., & Yu, J. (2022). A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in Sociology*, 7, 886498.

<https://doi.org/10.3389/fsoc.2022.886498>

Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2014). Diffusion of lexical change in social media. *PLoS ONE*, 9(11), e113114. <https://doi.org/10.1371/journal.pone.0113114>

Erbil, C., Özbilgin, M. F., & Mergen, A. (2025). Red pill leadership behaviours and discourse ethics. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-025-06194-4>

Evans, R. (2018, October 11). *From memes to infowars: How 75 fascist activists were “red-pilled”*. Bellingcat. <https://www.bellingcat.com/news/americas/2018/10/11/memes-infowars-75-fascist-activists-red-pilled/>

Farrell, T., Araque, O., Fernandez, M., & Alani, H. (2020). On the use of jargon and word embeddings to explore subculture within the reddit’s manosphere. *12th ACM Conference on Web Science*, 221–230. <https://doi.org/10.1145/3394231.3397912>

Farrell, T., Fernandez, M., Novotny, J., & Alani, H. (2019). Exploring misogyny across the manosphere in reddit. *Proceedings of the 10th ACM Conference on Web Science*, 87–96. <https://doi.org/10.1145/3292522.3326045>

Feltgen, Q., Fagard, B., & Nadal, J.-P. (2017). Frequency patterns of semantic change: Corpus-based evidence of a near-critical dynamics in language change. *Royal Society Open Science*, 4(11), 170830. <https://doi.org/10.1098/rsos.170830>

Fillmore, C. J., & Baker, C. (2012). A frames approach to semantic analysis. In B. Heine & H. Narrog (Eds), *The Oxford Handbook of Linguistic Analysis* (1st edn, pp. 313–340). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199544004.013.0013>

- Fish, S. (1982). *Is there a text in this class? - The Authority of Interpretive Communities*. Harvard University Press. <https://www.hup.harvard.edu/books/9780674467262>, see also <https://sites.english.unt.edu/~simpkins/Fish%20Acceptable.pdf>
- Geeraerts, D., & Cuyckens, H. (Eds). (2010). *The oxford handbook of cognitive linguistics* (1st edn). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199738632.001.0001>
- Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds), *Media Technologies* (pp. 167–194). The MIT Press. <https://doi.org/10.7551/mitpress/9780262525374.003.0009>
- Ging, D. (2019). Alphas, betas, and incels: Theorizing the masculinities of the manosphere. *Men and Masculinities*, 22(4), 638–657. <https://doi.org/10.1177/1097184X17706401>
- Gorgio, M. (2025, April 27). *What if neo took the blue pill*. Science Fiction Classics. <https://www.sciencefictionclassics.com/what-if-neo-took-the-blue-pill/>
- Gothard, K., Dewhurst, D. R., Minot, J. R., Adams, J. L., Danforth, C. M., & Dodds, P. S. (2021, May 25). *The incel lexicon: Deciphering the emergent cryptolect of a global misogynistic community*. arXiv.Org. <https://arxiv.org/abs/2105.12006v1>
- Gray, D., & Bernatzky, C. (2022). Red pill(Or pilling). In D. Della Porta, B. Klandermans, D. McAdam, & D. A. Snow (Eds), *The Wiley-Blackwell Encyclopedia of Social and Political Movements* (1st edn, pp. 1–3). Wiley. <https://doi.org/10.1002/9780470674871.wbespm621>
- Grundlingh, L. (2018). Memes as speech acts. *Social Semiotics*, 28(2), 147–168. <https://doi.org/10.1080/10350330.2017.1303020>
- Habermas, J. (2022). Reflections and hypotheses on a further structural transformation of the political public sphere. *Theory, Culture & Society*, 39(4), 145–171. <https://doi.org/10.1177/02632764221112341>

- Hagen, S., Tuters, M., & Wilson, H. (2020, May 27). *OILAB: Reactionary wokeness: How redpilling became a thing on reddit – oilab*. <https://oilab.eu/reactionary-wokeness-how-redpilling-became-a-thing-on-reddit/>
- Halle, M. (1962). Phonology in generative grammar. *WORD*, 18(1–3), 54–72.
<https://doi.org/10.1080/00437956.1962.11659765>
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In K. Erk & N. A. Smith (Eds), *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1489–1501). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/P16-1141>
- Haney-López, I. (2014). *Dog whistle politics: How coded racial appeals have reinvented racism and wrecked the middle class*. Oxford University Press. ISBN 10: 0199964270
<https://www.abebooks.de/9780199964277/Dog-Whistle-Politics-Coded-Racial-0199964270/plp?srsId=AfmBOoqLfR3ZxLeg4BOOAvyZbwrX-gN3FoZW5pZeRpbyxR5mOrFmQ9v1>
- Hasko, V. (2012). Qualitative corpus analysis. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (1st edn). Wiley. <https://doi.org/10.1002/9781405198431.wbeal0974>
- Heritage, F. (2023). *Incels and ideologies: Exploring how incels use language to construct gender and race*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-40184-8>
- Herring, S. (2007). A faceted classification scheme for computer-mediated discourse. *Language@Internet* ; 1 , 1, 4.
https://www.researchgate.net/publication/31590573_A_Faceted_Classification_Scheme_for_Computer-Mediated_Discourse
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In S. Barab, R. Kling, & J. H. Gray (Eds), *Designing for Virtual Communities in*

the Service of Learning (1st edn, pp. 338–376). Cambridge University Press.

<https://doi.org/10.1017/CBO9780511805080.016>

Herring, S. C., & Androutsopoulos, J. (2015). Computer-Mediated Discourse 2.0. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds), *The Handbook of Discourse Analysis* (1st edn, pp. 127–151). Wiley. <https://doi.org/10.1002/9781118584194.ch6>

Hoebanx, P. (2025). Red pill women: Heterosexual fantasies in misogynistic spaces. *Men and Masculinities*, 28(2), 177–196. <https://doi.org/10.1177/1097184X241286800>

Hurbik, K., Khan, N. M., & Khan, M. A. (2025). Language Change in the Digital Age: A Linguistics Analysis of Online Communication. *Social Sciences & Humanity Research Review*, 3(4). <https://doi.org/10.63468/sshrr.252>

Ignatow, G., & Robinson, L. (2017). Pierre Bourdieu: Theorizing the digital. *Information, Communication & Society*, 20(7), 950–966. <https://doi.org/10.1080/1369118X.2017.1301519>

Jaki, S., De Smedt, T., Gwózdź, M., Panchal, R., Rossa, A., & De Pauw, G. (2019). Online hatred of women in the Incels.me forum: Linguistic analysis and automatic detection. *Journal of Language Aggression and Conflict*, 7(2), 240–268. <https://doi.org/10.1075/jlac.00026.jak>

Jason, Z. (2015, April 28). Game of fear: The story behind gamergate. *Boston Magazine*. <https://www.bostonmagazine.com/news/2015/04/28/gamergate/>

Karjus, A., & Cuskley, C. (2024). Evolving linguistic divergence on polarizing social media. *Humanities and Social Sciences Communications*, 11(1), 422. <https://doi.org/10.1057/s41599-024-02922-9>

Kawaguchi, Y., Minegishi, M., & Viereck, W. (Eds). (2011). *Corpus-based analysis and diachronic linguistics* (Vol. 3). John Benjamins Publishing Company. <https://doi.org/10.1075/tufs.3>

Kozinets, R. V. (2020). *Netnography: The essential guide to qualitative social media research* (Third edition). SAGE. <https://uk.sagepub.com/en-gb/eur/netnography/book260905>

- Kress, G. R., & Leeuwen, T. van. (2020). *Reading images: The grammar of visual design* (Third edition). Routledge. ISBN 9780415672573
- Kurniasih, N. (2017). *Paralinguistics cues used and miscommunication on social media: Case study of students of communication science program universitas padjadjaran*.
<https://doi.org/10.31227/osf.io/fmp3n>
- Labov, W. (1995). William labov, principles of linguistic change. Volume i: Internal factors (Language in Society 20). Oxford: blackwell, 1994. Pp. Xix + 641. *Journal of Linguistics*, 31(2), 435–439. <https://doi.org/10.1017/S0022226700015693>
- Labov, W. (2001). Principles of linguistic change, volume ii: Social factors. *Social Factors*, 2.
https://www.researchgate.net/publication/247251889_Principles_of_Linguistic_Change_Volume_II_Social_Factors
- Laclau, E. (1996). *Emancipation(s)* (1. publ). Verso. ISBN 1844675769
https://books.google.de/books/about/Emancipation_s.html?id=QFwiHJP8mdUC&redir_esc=y
- Li, Y., & Siew, C. S. Q. (2022). Diachronic semantic change in language is constrained by how people use and learn language. *Memory & Cognition*, 50(6), 1284–1298.
<https://doi.org/10.3758/s13421-022-01331-0>
- Livingston, P. M. (2017). Twentieth-century philosophy. In *Routledge Encyclopedia of Philosophy* (1st edn). Routledge. <https://doi.org/10.4324/0123456789-DD3596-1>
- Lotman, J., & Clark, W. (2005). On the semiosphere. *Sign Systems Studies*, 33(1), 205–229.
<https://doi.org/10.12697/SSS.2005.33.1.09>
- Lucy, L., & Bamman, D. (2021). Characterizing english variation across social media communities with bert. *Transactions of the Association for Computational Linguistics*, 9, 538–556. https://doi.org/10.1162/tacl_a_00383

- Massanari, A. (2017). #Gamergate and The Fapping: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, 19(3), 329–346.
<https://doi.org/10.1177/1461444815608807>
- Matrix*. (1999). [Video recording]. Retrieved <https://www.themoviedb.org/movie/603-the-matrix>
- McCulloch, G. (2019). *Because internet: Understanding the new rules of language* (First Riverhead trade paperback edition). Riverhead Books. ISBN 0735210934
- McGlashan, M., & Krendel, A. (2024). Keywords of the manosphere. *International Journal of Corpus Linguistics*, 29(1), 87–115. <https://doi.org/10.1075/ijcl.22053.mcgl>
- Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and interpretation* (here reprint 2006). Blackwell. ISBN 0631222243
- Muhammed, H., & Arafah, B. (Eds). (2022). Narrativity and multimodal communication in the new media age. In *Proceedings of the Sixth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2022)* (Vol. 722, pp. 344–350). Atlantis Press SARL. https://doi.org/10.2991/978-2-494069-91-6_54
- Munn, L. (2019). Alt-right pipeline: Individual journeys to extremism online. *First Monday*.
<https://doi.org/10.5210/fm.v24i6.10108>
- Nagle, A. (2017). *Kill all normies: Online culture wars from 4chan and tumblr to trump and the alt-right*. John Hunt Pub. ISBN 1785355430
- Nevalainen, T., & Van Ostade, I. T.-B. (2006). Standardisation. In R. Hogg & D. Denison (Eds), *A History of the English Language* (1st edn, pp. 271–311). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511791154.006>
- Noll, A. (2025, August 30). *Taking the 'red pill': From meme to political ideology*. DW Deutsche Welle. Dw.Com. <https://www.dw.com/en/taking-the-red-pill-from-meme-to-political-ideology/a-73821029>

- O'Donnell, C., & Shor, E. (2022). "This is a political movement, friend": Why "incels" support violence. *The British Journal of Sociology*, 73(2), 336–351. <https://doi.org/10.1111/1468-4446.12923>
- Ogden, C. K., Richards, I. A., (1923), and Eco, U., Malinowski, B., & Crookshank, F. G. (1989 Reprint). *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism*. Harcourt Brace Jovanovich. ISBN 0-15-658446-8
- O'Halloran, K. L., E, M. K. L., Podlasov, A., & Tan, S. (2013). Multimodal digital semiotics: The interaction of language with other resources. *Text & Talk*, 33(4–5). <https://doi.org/10.1515/text-2013-0030>
- Okrent, A., & O'Neill, S. (2021). Blame the printing press. In A. Okrent & S. O'Neill, *Highly Irregular* (1st edn, pp. 116–146). Oxford University Press New York. <https://doi.org/10.1093/oso/9780197539408.003.0004>
- Phillips, W., & Milner, R. M. (2017). *The ambivalent Internet: Mischief, oddity, and antagonism online*. Polity. ISBN 1509501274
- Prem, E., & Krenn, B. (2024). On algorithmic content moderation. In H. Werthner, C. Ghezzi, J. Kramer, J. Nida-Rümelin, B. Nuseibeh, E. Prem, & A. Stanger (Eds), *Introduction to Digital Humanism* (pp. 481–493). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-45304-5_30
- Radulović, M. (2023). A review of methodologies and methods in linguistic research: Diachronic and synchronic approaches. *Facta Universitatis, Series: Linguistics and Literature*, 095. <https://doi.org/10.22190/FULL230330008R>
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3980–3990. <https://doi.org/10.18653/v1/D19-1410>

- Reisigl, M., & Wodak, R. (2017). The discourse-historical approach(Dha). *The Routledge Handbook of Critical Discourse Studies*.
https://www.researchgate.net/publication/251636976_The_Discourse-Historical_Approach_DHA
- Ribeiro, M. H., Blackburn, J., Bradlyn, B., Cristofaro, E. D., Stringhini, G., Long, S., Greenberg, S., & Zannettou, S. (2021). The evolution of the manosphere across the web. *Proceedings of the International AAAI Conference on Web and Social Media, 15*, 196–207.
<https://doi.org/10.1609/icwsm.v15i1.18053>
- Rorty, R. (1992). *The linguistic turn: Essays in philosophical method, with two retrospective essays*. the University of Chicago press. ISBN-10. 0226725693
- Rudolph, M., & Blei, D. (2018). Dynamic embeddings for language evolution. *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 1003–1011.
<https://doi.org/10.1145/3178876.3185999>
- Safire, W. (2008). *Safire's political dictionary*. Oxford university press.
<https://doi.org/https://books.google.de/books?id=jK-0NPoMiYoC>
- Saussure, F. de, Baskin, W., Meisel, P., Saussy, H., & Saussure, F. de. (2011). *Course in general linguistics*. Columbia University Press. ISBN 0231157266
- Schlechtweg, D., Zamora-Reina, F. D., Bravo-Marquez, F., & Arefyev, N. (2025). Sense through time: Diachronic word sense annotations for word sense induction and Lexical Semantic Change Detection. *Language Resources and Evaluation, 59*(2), 1431–1465.
<https://doi.org/10.1007/s10579-024-09771-7>
- Scrivens, R., & Gaudette, T. (2024). Online terrorism and violent extremism. In R. Scrivens & T. Gaudette, *Oxford Research Encyclopedia of Criminology and Criminal Justice*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264079.013.795>

- Shifman, L. (2014). *Memes in digital culture*. The MIT Press. ISBN 9780262525435
<https://mitpress.mit.edu/9780262525435/memes-in-digital-culture/>
- Shullenberger, G. (n.d.). Redpilling and the regime. *The New Atlantis*. Retrieved 23 April 2026, from <https://www.thenewatlantis.com/publications/redpilling-and-the-regime>
- Solopova, V., Popa-Wyatt, M., & Berškýtė, J. (2025). Exploring incel group dynamics: A computational study of hierarchy and group-boundary policing. *Journal of Computational Social Science*, 8(2), 27. <https://doi.org/10.1007/s42001-024-00357-5>
- Sontag, S. (2018). *Notes on 'camp'*. Penguin Books.
<https://www.penguin.co.uk/books/308595/notes-on-camp-by-sontag-susan/9780241339701>
 (Original work published 1964)
- Squires, L. (2010). Enregistering internet language. *Language in Society*, 39(4), 457–492.
<https://doi.org/10.1017/S0047404510000412>
- Steen, E., Yurechko, K., & Klug, D. (2023). You can (Not) say what you want: Using algospeak to contest and evade algorithmic content moderation on tiktok. *Social Media + Society*, 9(3), 20563051231194586. <https://doi.org/10.1177/20563051231194586>
- Stern, A. M. (2019). *Proud boys and the white ethnostate: How the alt-right is warping the american imagination*. Beacon Press.
- Sunstein, C. R. (2018). *#republic: Divided democracy in the age of social media*. Princeton University Press. <https://doi.org/10.2307/j.ctv8xnhtd>
- Tagliamonte, S. A., & Denis, D. (2008). Linguistic ruin? Lol! Instant messaging and teen language. *American Speech*, 83(1), 3–34. <https://doi.org/10.1215/00031283-2008-001>
- Tagliamonte, S., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language*, 85, 58–108. <https://doi.org/10.1353/lan.0.0084>
- Tahmasebi, N., Borin, L., Jatowt, A., Xu, Y., & Hengchen, S. (2021). *Computational approaches to semantic change*. Zenodo. <https://doi.org/10.5281/ZENODO.5040241>

- Taylor, A. (2021, December 21). *The Matrix's real-world legacy—From red pill incels to conspiracies and deepfakes*. <https://www.bbc.com/news/entertainment-arts-57572152>
- Thorne, S. L. (2003). Artifacts and cultures-of-use in intercultural communication. . *Language Learning & Technology* 7(2), 38–67.
https://www.academia.edu/1910662/Thorne_S_L_2003_Artifacts_and_Cultures_of_Use_in_Intercultural_Communication_Language_Learning_and_Technology_7_2_38_67
- Thurlow, C., & Mroczek, K. R. (Eds). (2011). *Digital discourse: Language in the new media*. Oxford University Press. <https://academic.oup.com/book/1880>
- Tiffany, K. (2021, April 13). The alt-right has lost control of ‘redpill’. *The Atlantic*.
<https://www.theatlantic.com/technology/archive/2021/04/red-pill-meme-alt-right-twitter/618577/>
- Traugott, E. C., & Dasher, R. B. (2001). *Regularity in semantic change*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511486500>
- Ullmann, s. (1962). *Semantics an introduction to the science of meaning*. Oxford blackwell. -
References—Scientific research publishing. (n.d.). Retrieved 6 April 2026, from
<https://www.scirp.org/reference/referencespapers?referenceid=743872>
- Van Dijk, T. A. (2008). *Discourse and power*. Palgrave Macmillan. Accessed via:
https://api.pageplace.de/preview/DT0400.9781137072993_A29487743/preview-9781137072993_A29487743.pdf
- Van Leeuwen, T. (2004). *Introducing social semiotics: An introductory textbook* (First Edition). Routledge. <https://www.routledge.com/Introducing-Social-Semiotics-An-Introductory-Textbook/Leeuwen/p/book/9780415249447> (Copyright 2005)
- Van Valkenburgh, S. P. (2021). Digesting the red pill: Masculinity and neoliberalism in the manosphere. *Men and Masculinities*, 24(1), 84–103.
<https://doi.org/10.1177/1097184X18816118>

- Van Veeren, E. (2026). "Taking the red pill: Conspiracy theories, gender, and the 'elusive epistemologies' of the manosphere. *International Feminist Journal of Politics*, 2026.
https://www.academia.edu/128807572/Taking_the_Red_Pill_Conspiracy_Theories_Gender_and_the_elusive_epistemologies_of_the_Manosphere
- Waite, T. (2023, April 11). *Based and redpilled: The FBI links internet slang to violent extremism*. Dazed. <https://www.dazeddigital.com/life-culture/article/59622/1/based-and-redpilled-the-fbi-links-internet-slang-to-violent-extremism-incel>
- Walkden, G., McCarley, G. H., Montero, R., Rolf, M., Einhaus, S., & Kauhanen, H. (2023). Sociolinguistic typology meets historical corpus linguistics. *Transactions of the Philological Society*, 121(3), 546–567. <https://doi.org/10.1111/1467-968X.12275>
- Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19(1), 52–90.
<https://doi.org/10.1177/009365092019001003>
- Waśniewska, M. (2020). The red pill, unicorns and white knights: Cultural symbolism and conceptual metaphor in the slang of online incel communities. In B. Lewandowska-Tomaszczyk (Ed.), *Cultural Conceptualizations in Language and Communication* (pp. 65–82). Springer International Publishing. https://doi.org/10.1007/978-3-030-42734-4_4
- Willaert, T., Van Eecke, P., Van Soest, J., & Beuls, K. (2021). A tool for tracking the propagation of words on Reddit. *Computational Communication Research*, 3(1), 117–132.
<https://doi.org/10.5117/CCR2021.1.005.WILL>
- Wittgenstein, L. (2014). *Tractatus logico-philosophicus: German and english*. Taylor and Francis. ISBN 9780203819807
- Würschinger, Q. (2021). Social networks of lexical innovation. Investigating the social dynamics of diffusion of neologisms on twitter. *Frontiers in Artificial Intelligence*, 4, 648583.
<https://doi.org/10.3389/frai.2021.648583>

Würschinger, Q., & McGillivray, B. (2025). Semantic change and socio-semantic variation: The case of COVID-related neologisms on Reddit. *Linguistics Vanguard*, 11(s3), 339–352.

<https://doi.org/10.1515/lingvan-2023-0106>

Zappavigna, M. (2015). Searchable talk: The linguistic functions of hashtags. *Social Semiotics*, 25(3), 274–291. <https://doi.org/10.1080/10350330.2014.996948>

Appendix

Subreddit	Total "redpill" mentions	First seen	Peak month	Peak mentions/mo	Semantic role	Status
r/The_Donald	51847	2016-01	2017-02	4185	Political awakening	Banned Jun 2020
r/TheRedPill	22491	2012-11	2014-09	533	Origin community	Quarantined Sep 2018
r/PurplePillDebate	15234	2013-08	2016-02	472	Contested debate	Active
r/asktrp	12876	2013-08	2014-05	292	TRP newcomer Q&A	Quarantined
r/MGTOW	12643	2013-02	2018-11	495	Male separatism	Banned Aug 2021
r/TheBluePill	8742	2013-02	2013-12	280	Parody / counter-discourse	Active
r/Braincels	4315	2017-12	2018-06	307	Blackpill crossover	Banned Sep 2019
r/4chan4trump	4152	2016-11	2017-08	1578	/pol/ Trump bridge	Banned
r/SubredditDrama	4102	2013-01	2014-01	155	Meta-commentary	Active
r/KotakuInAction	3847	2014-08	2014-11	124	GamerGate bridge	Active
r/milliondollarextreme	3612	2015-11	2018-05	287	Alt-right comedy	Banned Sep 2018
r/conspiracy	3548	2012-05	2017-02	209	Conspiracist awakening	Active
r/greatawakening	2733	2018-01	2018-07	705	QAnon & 'Great Awakening'	Banned Sep 2018
r/IncelTears	2314	2017-07	2018-11	103	Counter-discourse	Renamed
r/DebateAltRight	2187	2016-12	2019-08	138	Alt-right debate space	Banned
r/MensRights	2156	2012-04	2014-06	80	MRA overlap, pre-TRP	Active
r/altright	2042	2016-07	2016-12	477	Explicit white nationalism	Banned Feb 2017
r/RedPillWomen	1876	2013-06	2013-07	58	Gender inversion	Active

r/Incels	1743	2016-07	2017-10	211	Incel radicalisation	Banned Nov 2017
r/unpopularopinion	1876	2018-10	2020-03	153	Mainstream seepage	Active
r/exredpill	1687	2014-09	2020-01	87	Deradicalisation / exit	Active
r/niceguys	1654	2013-12	2017-04	44	Ironic mockery	Active
r/ChapoTrapHouse	1543	2017-05	2020-04	81	Left-wing use/mockery	Banned Jun 2020
r/PoliticalCompassMemes	1543	2019-08	2020-04	289	Canonised political slang	Active
r/marriedredpill	1098	2014-07	2015-08	54	Domestic TRP application	Active
r/IncelsWithoutHate	765	2017-11	2020-04	109	Moderate incel successor	Banned
r/IncelExit	124	2019-10	2020-04	32	Deradicalisation	Active

Data Set Audit

Word count percentiles:

p 1: 4 words
p 5: 8 words
p 10: 12 words
p 25: 25 words
p 50: 57 words
p 75: 124 words
p 90: 177 words
p 95: 185 words
p 99: 196 words

Very short (< 5 words) : 14,818 (1.53%)

Very long (> 500 words) : 0 (0.00%)

Top 30 authors by comment count:

	Redditor	count
0	AutoModerator (cleaned)	58311
1	[deleted] (cleaned)	57306
2	Grrrrreat (cleaned)	6046
3	SnapshillBot (cleaned)	4529
4	NeoreactionSafe	3492
5	GayLubeOil	2301
6	taiboworks	2006
7	exit sandman	1728
8	BluepillProfessor	1528
9	redpillschool	1420
10	wtknight	1384
11	Archwinger	1320
12	ProbablyBelievesIt	1219
13	sublimemongrel	1166
14	oldredder	1133
15	Soloandthewookiee	1075
16	Solondthewookiee	999
17	CisWhiteMaelstrom	980

18	Xemnas81	944
19	OfSpock	926
20	nomdplume	863
21	TheGreasyPole	844
22	disposable_pants	833
23	Whisper	795
24	Lenovo_Driver	769
25	GuitarsAreGettingOld	767
26	Atlas_B_Shruggin	754
27	bsutansalt	745
28	LillthOfBabylon	744
29	RedPillDetox	741

Top 1% of authors (1,638 users) account for 38.1% of comments
Probable bots (name contains bot/auto/moderator): 67,051

	count
Redditor	
AutoModerator	58311
SnapshillBot	4529
trpbot	394
strategos_autokrator	227
autowikibot	180
5th_Law_of_Robotics	145
RobotPartsCorp	127
PoopInMyBottom	118
la_sabotage	81
NotARussianBot1984	76
AutoDMC	72
sneakpeekbot	65
ActivatedJoeBot	61
bottleblank	60
RemindMeBot	58
HEDONISM_BOT	56
flakybottom	49
ERRORCORRECTB	41
OT	
RobotJINI	40
totes meta bot	39

Top 10 months by comment count:

year	month	count	year	month	str
2017	02	18621	2017	02	
2017	06	17007	2017	06	
2016	11	16752	2016	11	
2017	03	16734	2017	03	
2017	05	16362	2017	05	
2017	08	14812	2017	08	
2017	01	14643	2017	01	
2017	04	14588	2017	04	
2017	07	14583	2017	07	
2018	06	13721	2018	06	

Subreddits

	subreddit	count
0	PurplePillDebate	218139
1	The_Donald	173957
2	TheRedPill	149734

3	asktrp	79753
4	MGTOW	79276
5	TheBluePill	58612
6	exredpill	50939
7	RedPillWomen	30344
8	MensRights	23727
9	conspiracy	21740
10	marriedredpill	17077
11	KotakuInAction	9877
12	greatawakening	9345
13	niceguys	7528
14	IncelTears	6394
15	4chan4trump	6080
16	milliondollarextreme	4494
17	DebateAltRight	4491
18	Incels	4072
19	altright	3235
20	IncelsWithoutHate	2507
21	againstmensrights	1954
22	kotakuinaction2	1084
23	Braincels	940
24	IncelExit	896

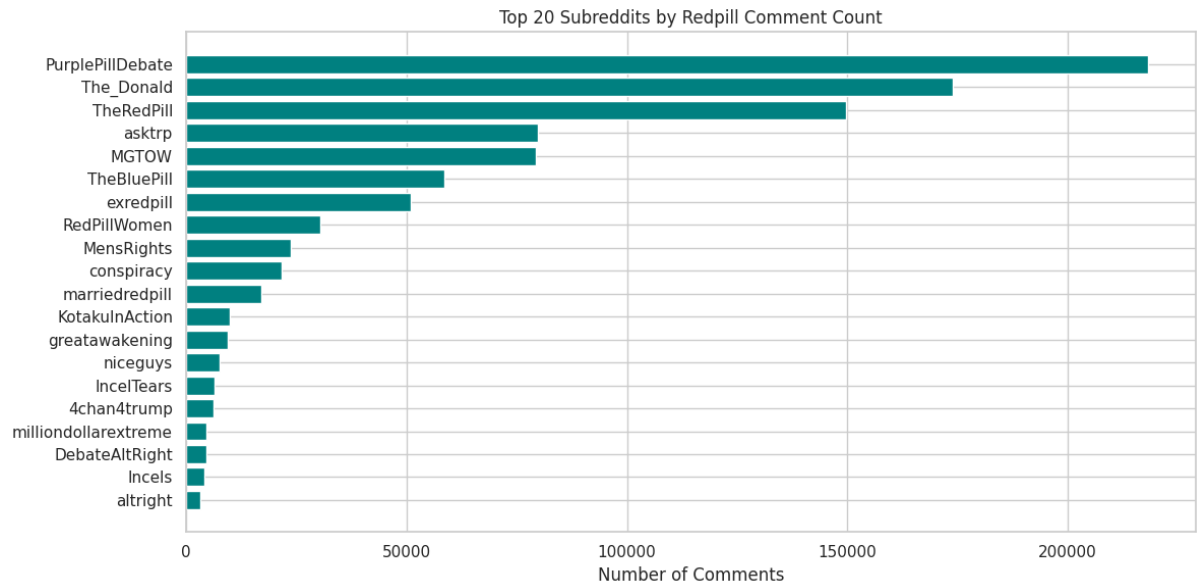
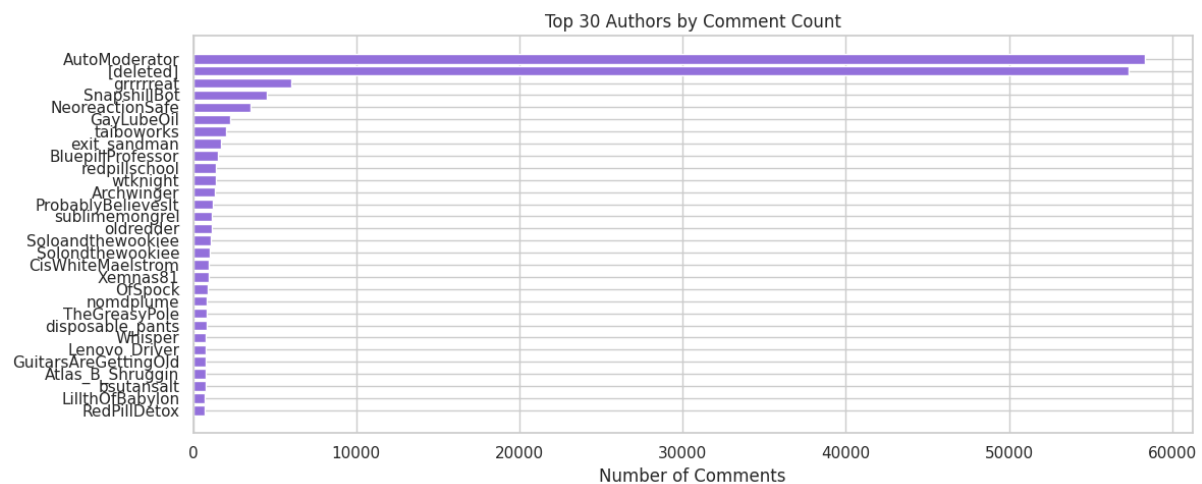
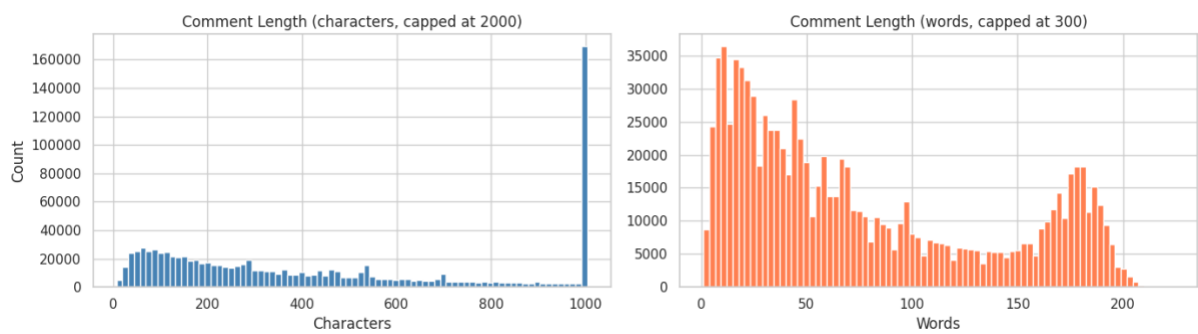
Top 3 subreddits = 56.1% of all comments
Top 5 subreddits = 72.5% of all comments
Top 10 subreddits = 91.7% of all comments
Top 20 subreddits = 99.2% of all comments
Unique texts appearing in >1 subreddit : 961

Top cross-posted texts:

	comment	n_subreddits
101466	/r/theredpill	14
163097	Based and redpilled	14
858970	based and redpilled	13
883208	r/TheRedPill	11
100708	/r/TheRedPill	11
163146	Based and redpilled.	11
565992	Red Pill	11
682087	The red pill	10
660607	The Red Pill	10
883756	r/theredpill	10
635583	Take the red pill	9
883521	r/marriedredpill	8
685729	The red pill.	8
774620	What is the red pill?	8
587417	Redpilled	8

Audit Summary

After removing deleted :	966,195
Unique authors :	163,842
Unique subreddits :	25
Date range :	2008-10-11 to 2024-12-31
Median comment length :	57 words
Probable bots :	67,051
Cross-posted texts :	961



"Redpill" Correlation Analysis

Pearson r + Spearman ρ across three observation windows (Reddit 2013–2024)

Period 1: 2013–2018

Observed data

Year	Comments	Subreddits	Redditors	Drift score
2013	23,549	9	4,706	0.000
2014	53,342	13	12,398	0.230
2015	66,525	15	16,009	0.380
2016	83,938	19	25,455	0.444
2017	153,113	22	42,672	0.445
2018	116,592	20	36,742	0.480

Correlation matrix

Pair	Pearson r		p	
Drift × Comment count	+0.819	0.046 *	+0.943	0.005 **
Drift × Subreddit diversity	+0.931	0.007 **	+0.943	0.005 **
Drift × Redditor diversity	+0.838	0.037 *	+0.943	0.005 **
Comment count × Subreddit div.	+0.954	0.003 **	+1.000	0.000 ***
Comment count × Redditor div.	+0.988	0.000 ***	+1.000	0.000 ***
Subreddit div. × Redditor div.	+0.968	0.002 **	+1.000	0.000 ***

Period 2: 2019–2024

Observed data

Year	Comments	Subreddits	Redditors	Drift score
2019	83,328	19	27,864	0.513
2020	50,611	19	15,956	0.510
2021	39,402	15	11,446	0.490
2022	38,021	14	9,831	0.466
2023	38,313	14	8,310	0.478
2024	36,208	14	7,300	0.479

Correlation matrix

Pair	Pearson r		p	
Drift × Comment count	+0.787	0.063 n.s.	+0.829	0.042 *
Drift × Subreddit diversity	+0.952	0.003 **	+0.926	0.008 **
Drift × Redditor diversity	+0.827	0.042 *	+0.771	0.072 n.s.
Comment count × Subreddit div.	+0.818	0.047 *	+0.926	0.008 **
Comment count × Redditor div.	+0.987	0.000 ***	+0.943	0.005 **
Subreddit div. × Redditor div.	+0.865	0.026 *	+0.926	0.008 **

Full period: 2013–2024 (n = 12 years)

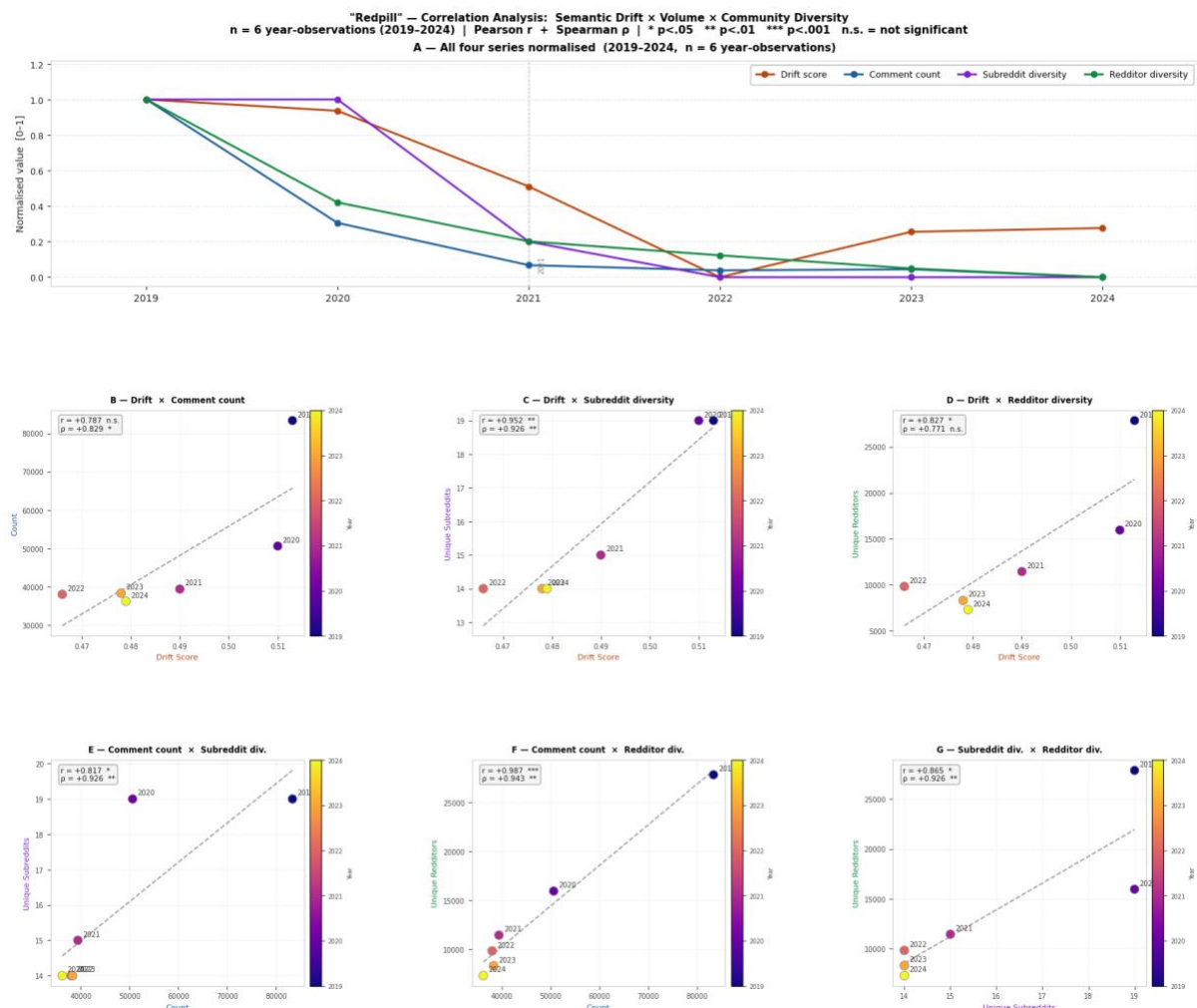
Observed data

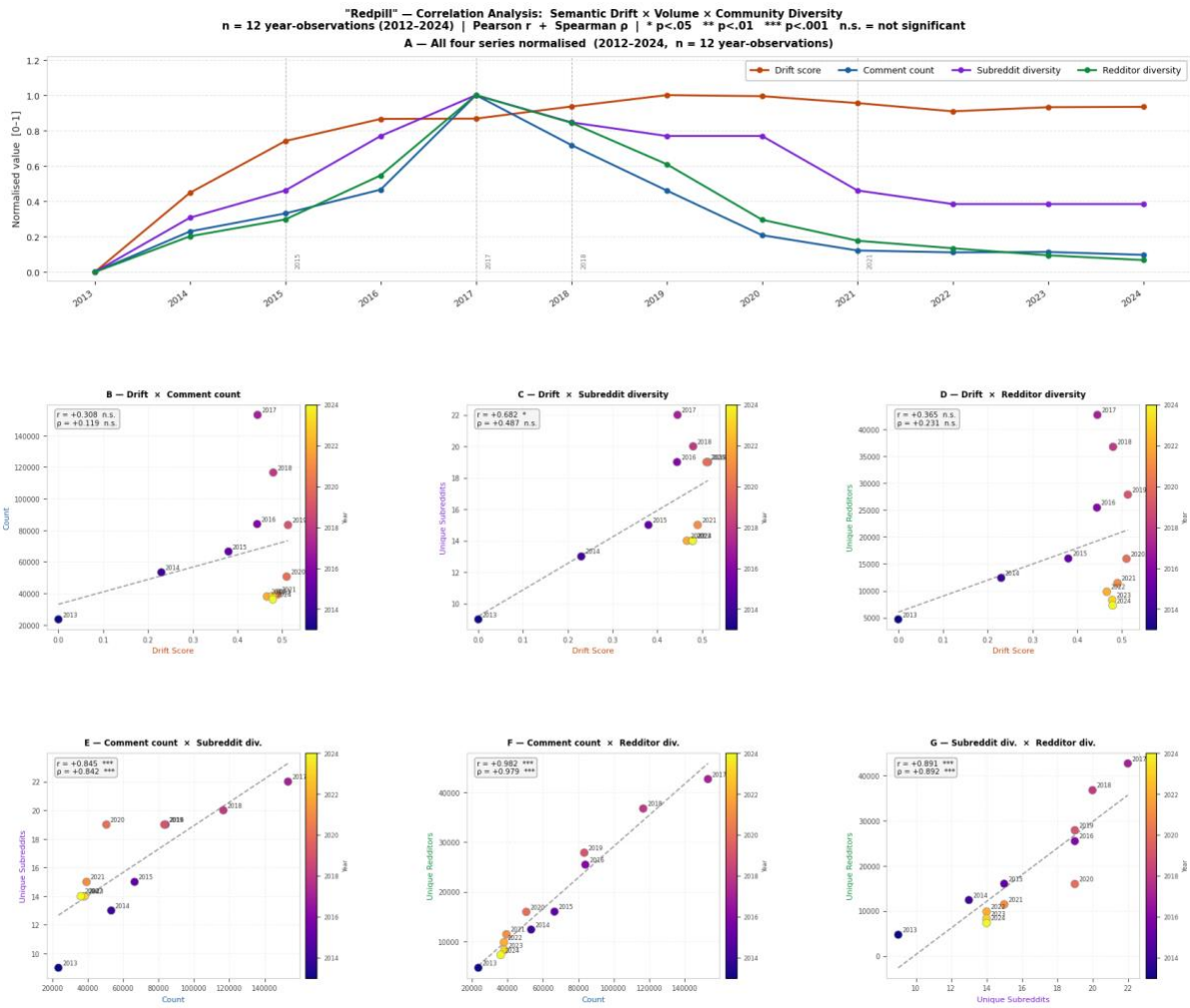
Year	Comments	Subreddits	Redditors	Drift score
2013	23,549	9	4,706	0.000
2014	53,342	13	12,398	0.230
2015	66,525	15	16,009	0.380
2016	83,938	19	25,455	0.444
2017	153,113	22	42,672	0.445
2018	116,592	20	36,742	0.480
2019	83,328	19	27,864	0.513
2020	50,611	19	15,956	0.510
2021	39,402	15	11,446	0.490
2022	38,021	14	9,831	0.466
2023	38,313	14	8,310	0.478
2024	36,208	14	7,300	0.479

Correlation matrix

Pair	Pearson r		p	
Drift × Comment count	+0.308	0.330 n.s.	+0.119	0.713 n.s.
Drift × Subreddit diversity	+0.682	0.015 *	+0.487	0.109 n.s.
Drift × Redditor diversity	+0.365	0.243 n.s.	+0.231	0.471 n.s.
Comment count × Subreddit div.	+0.845	0.001 ***	+0.842	0.001 ***
Comment count × Redditor div.	+0.983	0.000 ***	+0.979	0.000 ***
Subreddit div. × Redditor div.	+0.891	0.000 ***	+0.892	0.000 ***

Significance: *** p < .001 ** p < .01 * p < .05 n.s. not significant · Pearson r = linear · Spearman ρ = rank-order





Abstract

This thesis traces the semantic trajectory of *redpilled* across Reddit from 2012 to December 2024, developing a mixed-method framework that combines diachronic distributional semantics with semiotic theory and interpretive discourse analysis. While diachronic word embeddings are technically established and manosphere discourse extensively documented qualitatively, no study has traced a single ideologically charged, digitally native token across an ideologically diverse ecosystem along both the temporal and the cross-community axis simultaneously.

The study draws on approximately 966,000 comments from 25 subreddits reconstructed from the Pushshift archive. Thirteen annual skip-gram Word2Vec models with a five-token context window are trained on a corpus in which all morphological variants of the target are normalized to a single placeholder token and pill-family lexemes excluded from the context window, preventing the embedding space from anchoring on intra-community lexical convention. The annual spaces are aligned through orthogonal Procrustes rotation to a 2012–2014 *r/TheRedPill* baseline, and drift is operationalized through three cosine-distance measures: cumulative displacement, year-on-year incremental change, and a volume-weighted global statistic. Synchronic variation is captured through subreddit-level nearest-neighbor extraction and zero-shot sentence-embedding classification (all-MiniLM-L6-v2) against theoretically derived field descriptors, requiring no annotated training data. Community Semantic Profiles situate each subreddit barycentrically across semantic fields; drift is additionally fitted to a logistic S-curve (asy. $L = 0.48$) and correlated with comment volume and actor diversity. Computational results are interleaved with netnographic close reading and a full lexicographic profile.

Cumulative drift rises sharply between 2013 and 2016 before stabilizing in the 0.44–0.52 range, with incremental peaks in 2013, 2016, and 2020 coinciding with Gamergate, the Trump campaign, and the mainstreaming of QAnon; the corpus further registers exogenous shocks such as the manosphere's mass-violence episodes and its participating actors as well as sustained metalinguistic deliberation over the term's meaning. Rather than evolving through successive stages, *redpilled* accumulates as a stratigraphy of overlapping registers (manosphere, MAGA, QAnon, incel, contested debate, counter-discourse), none fully displacing its predecessors. The thesis therefore advances a quantitatively grounded account of unlimited semiosis: a process in which each reappropriation deposits a further stratum without canceling those beneath it. As the term migrates outward, its ideological potency decouples from its semantic charge: it remains culturally legible while becoming ideologically plural, approaching the condition of an open signifier.

Platform moderation drives rather than brakes this process. Successive bans did not eliminate the term but dissolved the interpretive communities that had authorized its meaning, leaving a sign that survives lexically while losing ideological precision. The thesis argues that content moderation functions as a form of semantic governance whose semiotic consequences outlast the communities it removes.

Keywords: *computer linguistics; semantic change; diachronic word embeddings; distributional semantics; semiotics; unlimited semiosis; computational discourse analysis; natural language processing; content moderation; manosphere; Reddit; NLP*