

CULTURAL NORMATIVITY IN THE ARAB-ISLAMIC WORLD

TRACING TIGHTNESS ACROSS 1,400 YEARS OF TEXTUAL TRADITION

Emna BEN AMEUR - Safia LAMRI

Supervised by Nicolas BAUMARD & Valentin THOUZEAU

Institut Jean Nicod – École Normale Supérieure (PSL)

June 2025

This study was conducted as part of a UROP internship at the Institut Jean Nicod (ENS – PSL).

Contents

1	Introduction	3
2	Data	4
3	Methodology	4
4	Results	5
5	Perspectives	7

1 Introduction

Are Arab-Islamic societies inherently governed by strict moral and social norms?

This question, central to both internal discourses and external representations, evokes the notion of *tightness* — the degree to which behavioral norms are explicit, enforced, and sanctioned within a culture (Gelfand et al., 2011).

While the perception of the Arab-Muslim world as “tight” is widespread, it remains unclear whether this tightness is a constant historical feature, the result of a slow accumulation, or the consequence of a specific turning point. Has the normativity of texts remained stable, intensified progressively, or undergone sharp transformations at key historical moments?

A major hypothesis in this regard comes from economic historian Eric Chaney (2015), who identifies a cultural shift in the 12th–13th centuries that altered the production of knowledge and religious authority. This raised a broader question for us: can such macro-historical dynamics be observed in the internal characteristics of written texts? Do periods of cultural openness and prosperity correspond to more permissive textual norms, and do times of political or social instability correlate with a hardening of discourse?

To investigate this, we draw on the concept of *cognitive fossils* — linguistic traces of cultural values and psychological orientations embedded in texts — as formulated in recent work on historical psychology (Baumard et al., 2023). Rather than relying on institutional records or external events, we aim to track normativity through language itself, over more than 1,400 years of Arabic scholarly production.

Our goal is to test whether—and how—the expression of normative concepts evolves across genres and centuries, and whether this evolution echoes historical transformations or instead unfolds along a more autonomous trajectory within the Arab-Islamic textual tradition.

2 Data

Our dataset is drawn from the *Open Islamicate Texts Initiative* (OpenITI), a large-scale corpus of over 8,000 Arabic texts collected from the Shamela digital library and curated by academic researchers. This corpus spans more than 1,400 years of literary production and includes a wide variety of genres and authors.

For this study, we selected a subset of 3,741 texts, which are the most consistently preprocessed and reliably dated. We focused on three discursive genres with strong normative intent:

- **LEG (Legal)**: treatises on Islamic jurisprudence (*fiqh*) and legal reasoning.
- **TRA (Tradition)**: compilations of *ḥadīṭ* (prophetic traditions).
- **EXE (Exegesis)**: *tafsīr* texts interpreting the Qur’ān.

This selection covers a wide chronological range — from the 7th to the 20th century — while maintaining a stable thematic focus on moral, religious, and legal normativity.

Each file includes metadata such as the author’s name, title, and death date (in Hijri), which serves as a proxy for the text’s time of composition.

Basic preprocessing steps were applied to remove corrupted files, incomplete metadata, and duplicates. No linguistic annotation or normalization was performed at this stage.

3 Methodology

To trace long-term shifts in discursive normativity, we apply a lexical approach designed to detect semantically coherent markers of tightness and looseness within classical Arabic texts. Our method builds on the *Bag-of-Concepts* framework (Alahmadi et al., 2014), which enables the identification of abstract values beyond surface word forms — a crucial feature when working with morphologically rich languages such as Arabic.

As a first step, we applied a standardized lemmatization procedure to each text, reducing inflected words to their base forms and unifying orthographic variants. This preprocessing allowed us to map diverse lexical items to stable conceptual categories, ensuring consistency across time periods and authors.

We then developed two targeted lexicons representing opposing normative orientations:

- **Tightness-related concepts**, expressing rigidity, constraint, and normative enforcement. Examples include: *ḥukm* (judgment), *ilzām* (obligation), *ḥaram* (prohibition), *'iqāb* (punishment).
- **Looseness-related concepts**, associated with openness, flexibility, and individual latitude. Examples include: *murūna* (flexibility), *ḥurriyya* (freedom), *ibdā'* (creativity), *tasāmuḥ* (tolerance).

For each text, we calculated a **tightness score** defined as the proportion of tightness-related lemmas among all normative lemmas (tight + loose). This score ranges from 0 (maximally loose) to 1 (maximally tight), and captures the relative weight of rigid vs. permissive concepts within a given document.

To model historical trends, texts were grouped by century based on the author's death date, converted from Hijri to Gregorian. Aggregating the scores across periods allows us to observe diachronic evolutions in normative discourse — both within each genre and across the dataset as a whole.

4 Results

Evolution of Tightness Scores

The following figure illustrates the diachronic distribution of tightness scores across the three genres studied — exegesis (EXE), legal (LEG), and hadith (TRA). Each dot represents an individual text, positioned according to the author's death date (converted to the Gregorian calendar). Genre-specific linear regression lines are superimposed to capture overall trends.

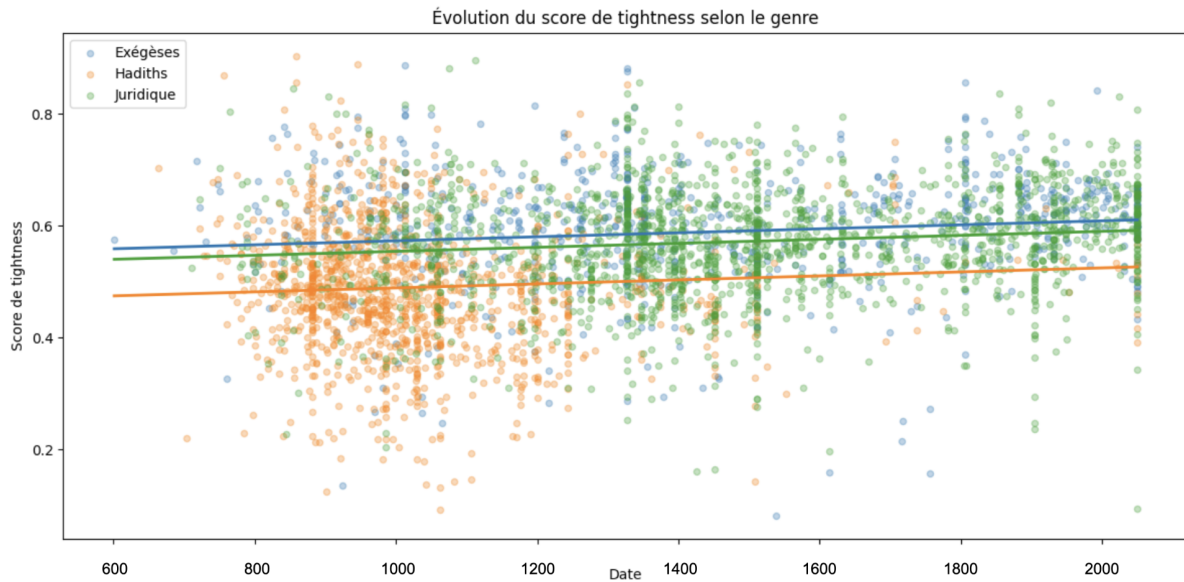


Figure 1: Diachronic distribution of tightness scores by genre (EXE, LEG, TRA).

Key Findings

A strong genre effect. Clear differences emerge across genres:

- **EXE** (Exegesis) texts display the highest tightness scores on average, reflecting dense normative expression and interpretive authority.
- **LEG** (Legal) texts exhibit intermediate but stable scores, aligned with their prescriptive function.
- **TRA** (Hadith) texts show the lowest tightness levels, likely due to their fragmentary and narrative structure.

These distinctions confirm that the lexical tightness score captures meaningful differences in normative density between genres.

A continuous rise in normativity. Contrary to the hypothesis of a sharp normative shift in the 12th–13th centuries, our data reveal no visible rupture in the tightness trend. Instead, we observe a gradual and linear increase in normative density across the centuries, beginning in the earliest periods of textual production.

This trend points to a slow but consistent accumulation of normative expression over time — potentially driven by the increasing formalization of religious and legal knowledge.

These findings support the hypothesis that Arab-Islamic societies may have become progressively “tighter” over time, in parallel with the institutionalization and codification of normative discourses. Rather than responding to punctual historical events, discursive normativity appears to follow a trajectory of cultural consolidation and textual stabilization.

Model validation. We tested several models to capture the temporal evolution of tightness scores. The linear model proved to be the most significant and parsimonious, offering a robust fit without overfitting. Quadratic models and changepoint detection methods failed to reveal any statistically significant rupture.

This reinforces the interpretation of normativity as a gradual sedimentation process — rather than the effect of discrete historical transformations — and highlights the utility of tightness as a lexical proxy for tracking long-term cultural dynamics.

5 Perspectives

While our analysis reveals a gradual increase in normative density across centuries and genres, it also opens new avenues for exploration — both within and beyond the Arab-Islamic textual tradition.

Beyond High-Normative Genres Our analysis focused on three highly prescriptive genres (law, exegesis, and tradition). Future work could extend this investigation to less normative or more narrative genres — such as poetry, biography, or historical chronicles — to examine whether the linguistic expression of normativity differs significantly in texts with other epistemic or rhetorical aims.

Comparative Corpora A comparative study involving other textual traditions — such as Persian, Latin, or Hebrew — could help identify whether the dynamics of normative accumulation we observed are specific to the Arab-Islamic tradition or reflect broader patterns in the historical evolution of cultural tightness.

Reception and Social Practices Finally, understanding normativity requires looking beyond textual production. How were these prescriptive texts received, interpreted, taught, negotiated or bypassed in real social contexts and actual scholarly or legal practice? Combining our lexical analysis with ethnographic inquiry or interviews with contemporary

scholars in *fiqh*, *tafsīr*, and *ḥadīth* studies could illuminate the lived dimension of normative transmission — where rules are not only written, but enacted, challenged, or reimagined.

Taken together, these extensions would allow for a richer account of how cultures articulate and transmit moral constraints — and of how tightness, as a psychological and linguistic orientation, becomes sedimented in discourse over time.

References

- [1] Gelfand, M., Jackson, J. C., De, S., & Fox, A. (2023). The loosening of American culture over 200 years is associated with a creativity–order trade-off. *Nature Human Behaviour*. [doi/URL if available].
- [2] Chaney E. (2015). Religion and the Rise and Fall of Islamic Science. In E. Berman & C. W. Makdisi (Eds.), *Scholars, Saints, and Sufis: Intellectual and Religious Life in the Islamic World*. [Publisher if known].
- [3] Alahmadi A., Joorabchi A., & Mahdi A. E. (2014). Arabic Text Classification using Bag-of-Concepts Representation. Electronic and Computer Engineering Department, University of Limerick.
- [4] Baumard N., Safra L., Martins M., & Chevallier C. (2023). Cognitive fossils: Using cultural artifacts to reconstruct psychological changes throughout history. [Preprint or journal name if available].
- [5] Romanov M. (2017). Observations of a Medieval Quantitative Historian? [Blog post or online article, add URL if needed].