International Journal of Law in Context (2026), **22**, 340–361
doi:10.1017/S1744552326100469

ARTICLE

Consensus and legitimation in global AI regulations: a sociosemiotic perspective

Jiaxuan Qiu and Le Cheng

Guanghua Law School, Zhejiang Univeristy, Hangzhou, Zhejiang, China
Corresponding author: Le Cheng; Email: chengle163@hotmail.com

(Received 15 July 2025; revised 18 November 2025; accepted 18 February 2026; first published online 06 April 2026)

Abstract
This paper adopts a sociosemiotic perspective to examine how normative consensus and legitimacy are constructed in global artificial intelligence (AI) governance discourse. Drawing on a corpus of forty-seven international normative documents, the study identifies an emerging cross-textual consensus around three core principles – Safety, Human-centric and Fairness – and analyses how these are semiotically encoded. The findings reveal tensions between state and non-state actors, and between semiotic agreement and practical implementation. For instance, 'Safety' is often framed through securitisation discourse, while 'Human-centric' is increasingly grounded in international human rights frameworks. The study further shows that discursive strategies such as nominalisation help establish surface-level consensus but introduce ambiguity that undermines enforceability. By conceptualising governance texts as dynamic semiotic systems, this research moves beyond the hard law–soft law dichotomy, revealing global AI regulation as a contested arena of meaning-making. It offers a theoretical basis for advancing more inclusive and operational governance models.

Keywords: artificial intelligence regulation; sociosemiotics; global governance; soft law; consensus-building; human rights framework

1. Introduction

The rise of generative artificial intelligence (AI) and its transformative impact pose novel legal challenges, ethical dilemmas and even existential risks for global stakeholders (Dwivedi *et al.* 2021; Fui-Hoon Nah *et al.* 2023; Wirtz *et al.* 2019). Governments, international organisations, multinational corporations and research institutions are continuously issuing diverse regulations, principles, frameworks and strategies in pursuit of global oversight of AI. However, divergent national interests and varied governance stances (Clarke 2019; Smuha 2021) have led to previous international normative outputs focusing predominantly on ethical guidelines, policy documents and technical standards, which typically lack binding consensus (Schiff *et al.* 2020; Schmitt 2022; Villarino 2023).

The current global AI regulatory landscape must be understood within the context of the global AI governance regime (Veale *et al.* 2023). AI's cross-border externalities (Jonas *et al.* 2023) render unilateral state efforts insufficient, prompting calls for broader stakeholder participation in shaping AI regulatory frameworks. Indeed, the very concept of 'the Global' is not a pre-existing container for state action, but is itself constructed through various legal and political practices that

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

increasingly transcend territorial boundaries (Rajkovic 2025). Terminologically, 'governance' extends beyond 'government', encompassing not only traditional international bodies but also overlapping, interlinked regional institutions and the growing involvement of non-state actors (Levi-Faur 2012). The international AI regulation regime reflects the paradigm of global governance. Current global AI regulations lack a clear central authority or hierarchy (Erman and Furendal 2022), instead consisting of fragmented, interlocking arrangements built on a patchwork of old and new legal frameworks across jurisdictions, including transnational private standards, best practice initiatives and normative declarations/principles from international organisations (Green and Auld 2017; Marchant and Allenby 2017). This complex web of norms, produced and enforced by a variety of actors, is a prime example of what the sociologist of law Eugen Ehrlich famously called 'living law' (*lebendes Recht*) – the law that actually dominates social life, often independently of formal state legislation. From this perspective, much of the AI regulatory landscape can be seen as a form of non-state law, emerging from the 'inner order' of social associations rather than from sovereign command (Hertogh 2008a; Hertogh 2008b; Nelken 2008).

Global norms for regulating AI constitute an evolving and dynamic corpus, largely due to their foundation in international 'soft law'. Despite critiques of soft law from both positivist and critical legal scholars (Kennedy 1980; Klabbers 1998; Koskenniemi 2006; Weil 1983), its rise in technology governance, coupled with the slow pace of hard law, poses a significant challenge to traditional state-centric international legal studies. Declining extending research to these emerging social realities risks ceding the study of global AI regulation to disciplines like political science, IR or management. Recognising soft law's potential virtues, modern legal positivists, rationalists and constructivists have integrated its study into their frameworks. For instance, Kratochwil (1991), Onuf (1985) and Günther (2008) combine legal positivism with sociological approaches, proposing discourse-based methods to understand soft law and international legal systems. Onuf (1985) applied speech act theory to show that international order, including soft law like UN resolutions, functions as a legal order. Günther (2008), building on Luhmann (1995), views international law as a communicative system and a certain form of discourse, arguing that norms become law by adhering to normative principles, thereby encompassing soft law.

How can we characterise this dynamic and evolving international AI governance regime primarily constituted by soft law? Drawing linguistic inspiration from scholars like Onuf and Günther, this paper proposes applying sociosemiotics as a methodological lens. Semiotics has been innovatively applied to legal studies (Balkin 1990; Cheng and Danesi 2019; Cheng and Sin 2008; Jackson 1985; Kennedy 1991; Kevelson 1988). Social semiotics specifically focuses on the relationship between signifiers and signified (the meaning-making process) within social behaviour and interaction (Cobley 2010). Within this framework, both texts and social systems are understood as sign systems (Halliday 2014), highlighting sociosemiotics' role as a meta-discipline for delineating sociocultural phenomena and sign systems (Cobley and Randviir 2009). This perspective provides an analytical anchor point for examining the fluid international legal governance of AI. By viewing AI normative texts, the AI governance regime and international society as multilayered sign systems, we can examine the semiotic interplay across discursive, regulatory and social dimensions. This approach enables the deconstruction and reconstruction of AI international regulation, identifying key characteristics of the current regime, revealing the underlying semiotic processes and socio-ideological factors, and helping refine our ethical and legal intuitions on this issue (Balkin 1990).

2. Theoretical framework

This study focuses on AI regulation and first clarifies the characteristics of international AI regulation to demonstrate its alignment with the sociosemiotic framework. A defining feature is its fragmentation, with norms dispersed across distinct international regimes. International regimes constitute patterned behaviour (Haggard and Simmons 1987), emphasising the normative

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

dimension of international politics and reflecting the interplay between international politics and law. They comprise issue areas, actors/organisations, norms and decision-making procedures (Krasner 1982). Currently, norms governing AI are widely distributed across regimes like international human rights, telecommunications and trade. Examples include using human rights principles for AI algorithmic accountability frameworks (McGregor *et al.* 2019), regulating cross-border data flows for AI development via telecommunications law (Taylor 2019), and addressing AI's legal complexities within international trade law (Igbinenikaro and Adewusi 2024). A dedicated international AI regime remains nascent.

A second characteristic of international AI regulation is its predominance of soft law and lack of hard law. Key actors include inter-governmental organisations (UN, G7, G20, BRICS Alliance, ASEAN) alongside numerous non-governmental and multi-stakeholder bodies (ITU, GPAI, IEEE, ITO). The norms they produce, being primarily non-binding, fit Ehrlich's classic distinction between the state's formal 'norms for decision' (*Entscheidungsnormen*) and the society's 'living law' which comprises the actual rules of conduct that people follow (Ehrlich and Ziegert 2017). Norms produced are primarily soft law or even 'ultra-soft law', with inter-governmental instruments achieving the broadest consensus, exemplified by UNESCO's Recommendation on the Ethics of Artificial Intelligence, unanimously adopted by all 193 member states. 'Ultra-soft law' encompasses self-regulatory industry standards; their technical expertise, filling international legal gaps, may generate strong appeal for compliance. Similar to the Tallinn Manual's significance for cyber law, standards from bodies like ISO/IEC hold considerable weight and may frame future AI regulatory agendas through path dependency (Villarino 2023).

A third distinctive feature is the plurality of regulatory subjects. This manifests in two ways: first, the diversification of international legal subjects and objects, where individuals and private firms now bear international legal rights, duties and liabilities under indirectly applicable regimes (e.g. international human rights law requiring technical accountability for AI systems (McGregor *et al.* 2019)). Second, diverse actors increasingly participate in the norm-making process of international AI regulation. Legal pluralism and liberalism in IR provide theoretical grounding for extending soft law authorship (Burley 1993; Merry 1988; Teubner 1991). Globalised AI technology intensifies risk society dynamics (McGrew 2004), necessitating regulation by expert committees and epistemic communities for complex, technical global issues. Their 'ultra-soft law', due to its expertise, often provides de facto guidance (Evans 2014, p. 411), acting as model rules that shape subsequent regulatory agendas. Furthermore, individuals, private firms and NGOs are increasingly visible in international forums/processes. For instance, at the inaugural 2023 AI Safety Summit in London, alongside governments, executives and researchers from leading tech firms like Google and Meta participated, contributing to *The Bletchley Declaration*.

The fragmentation, soft law dominance and plurality of actors in AI regulation transcend the public–private law dichotomy, exhibiting characteristics of social legislation. Current AI-specific regulatory instruments do not conform to any source of international law listed in Article 38(1) of the ICJ Statute (reflecting the sources under the Vienna Convention on the Law of Treaties). As the object of this study, these phenomena necessitate a theoretical framework that: (1) explains AI regulatory signs across multiple regimes; (2) accounts for the dynamic evolution of international AI regulation, with explanatory power for ultra-soft law, soft law and hard law; and (3) interprets the roles of contested legal actors interacting at the international level.

To address the aforementioned research requirements, this paper employs a sociosemiotic framework. Ferdinand de Saussure, the founder of semiotics, divided the structure of signs into *signifier* (the form carrying meaning) and *signified* (the concept or meaning) (De Saussure 2004). By limiting his research to linguistics, Saussure characterised the signifier–signified relationship as 'arbitrary', while the social, political and cultural factors influencing both were termed extra-semiotic phenomena – precisely the focus of sociosemiotics. Sociosemiotics prioritises analysing the structure of the signifier, its a priori meaning as a sign, and how the union of signifier and signified encodes social reality. This framework treats all semiotic processes as social processes:

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

social processes concern the definition of social actors, relations, structures and procedures, whereas semiotic processes serve as the means through which these definitions are examined, reaffirmed and transformed (Hodge and Kress 1988).

From a sociosemiotic perspective, all legal terms function as signs, while the international AI governance regime itself constitutes a dynamic semiotic system; the context of semiosis serves as the primary guide for interpretation (Cobley and Randviir 2009, p. 2). Diverse forms of AI regulation – including declarations, standards and recommendations – are codified as textual artefacts. As contextualised instantiations of social meaning, these texts manifest the relationship between semantic structures and social context, a relationship subject to sociosemiotic representation (Hymes 1971). Halliday's tripartite variables explain the semiotic structure of context: the field (the nature of social activity), tenor (participant relationships) and mode (medium of meaning exchange) (Halliday 2014, p. 192). Such tripartite division might cause some overlapping effects, but the configuration of the three elements into register integrates communication with social factors. Consequently, the sociosemiotic framework transcends the fragmentation of international regimes, enabling systematic deconstruction of all regimes through semiotic variables.

Second, sociosemiotics provides a robust framework for interpreting discursive dynamics in international AI governance. Inherently concerned with change and conflict, sociosemiotics focuses on explaining transformations, discontinuities, tensions and contradictions within human interactions and social processes. Sociosemiotics deciphers social constructs – including class stratification, power systems and hierarchies – through its diachronic and synchronic dimensions. When reconstructing the discourse, sociosemiotics reveals how social reality is shaped, constrained and modified (Berger and Luckmann 2016). The international governance of AI represents a convergence of ethical imperatives, political interests and law-making endeavours. Examining current regulatory discourse through a semiotic lens transcends superficial linguistic manoeuvers, thereby exposing the mutual permeability between morality, politics and law.

A sociosemiotics perspective is adaptive to the plurality of actors and multi-stakeholder dynamics in international AI governance. This framework treats motives (driving sign usage) and interests (guiding sign interpretation) as core factors in meaning production (Hodge and Kress 1988). Within Halliday's contextual variables, tenor addresses interpersonal relations, reflecting the 'participatory' function of language that encodes status and role hierarchies among actors (Halliday 2014, p. 125). Crucially, the concept of 'participants' here denotes not individuals but social actors embedded in institutional contexts and value systems. Through their continuous deconstruction and reconstruction of signs, semantic fluidity emerges. This is manifested in intertextual dynamics across regulatory documents: when successive texts invoke similar terms, the evolving referent of signs progressively re-shapes legal concepts in AI governance.

Viewing international AI regulation as a semiotic structure integrates regulatory texts with discursive processes (Hodge and Kress 1988, p. 262). As Figure 1 illustrates, regulatory texts function as signifiers of governance, where semantic/lexical relations encode critical traces of semiosis. Regulatory discourse operates as the mediating layer between text and social context, simultaneously constituting textual elements and social practices (Fairclough 2003, p. 45). This discursive dimension, being inherently semiotic, allows critical discourse analysis to dissect legitimisation mechanisms linguistically (Vaara *et al.* 2006; Van Leeuwen and Wodak 1999). The social context of AI governance, structured semiotically through field, tenor and mode variables, forms an integrated framework linking situational, discursive and textual strata.

3. Data and methods

To ground the abstract principles of social semiotics in a concrete analysis of legal texts, this study utilises a corpus-assisted methodology. While previous research on AI soft law has compiled valuable datasets (Fjeld *et al.* 2020; Jobin *et al.* 2019; Schmitt 2022), the reliance on purely qualitative methods presents challenges of scale and potential interpretive subjectivity. A corpus-

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

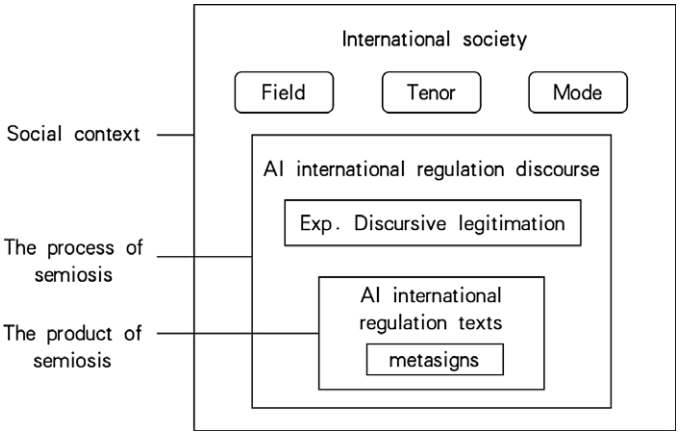


Figure 1. The semiotic structure of international AI regulation.

based approach addresses these challenges by enabling the systematic, large-scale analysis of linguistic patterns as an empirical foundation for qualitative interpretation. As an established method in legal linguistics (Cheng *et al.* 2019; Goźdź-Roszkowski 2021; Solan and Gales 2017; Vogel *et al.* 2018), corpus analysis facilitates the objective identification of keywords, collocational patterns and semantic trends, thereby revealing the underlying structure of a discourse before a more nuanced reading.

The first step in this research was to construct a specialised target corpus. Recognising that the explosive growth of AI soft law makes comprehensive collection nearly impossible, the collection process was based on the Global Partners Digital 'Navigating the Global AI Governance Landscape' report.[1] From their extensive catalogue of initiatives, formal normative documents were identified and then filtered to include only those that explicitly link specific actors to established norms; the set was also updated with key regulations released through 2024. This curated process yielded a target corpus of forty-seven documents representing a snapshot of the core international AI regime, with each document featuring a clear problem domain and institutional anchor (Appendix 1).

The distribution of drafting institution types is shown in Figure 2. To ensure the comparability of corpus length across document types in the discursive legitimation analysis, excessively long documents such as the EU AI Act were not included in the target corpus. However, the target corpus already includes other EU-issued documents, such as international conventions and advisory soft law files containing similar principles, to ensure the representation of the EU's perspective. Overall, the target corpus is primarily composed of soft law (such as declarations and recommendations) and ultra-soft law (such as industry standards). Hard law documents mainly include two AI conventions and representative domestic AI laws from influential governments that have binding force and international impact.

The analytical procedure began with preparing the documents for computational analysis. The forty-seven files were converted to .txt format and imported into AntConc (v. 4.3.1). For comparative analysis, the Corpus of Contemporary American English (COCA) was loaded as a reference corpus, and the Harbin Institute of Technology (HIT) stop word list was applied to exclude function words. Table 1 presents the details of two corpora. The analysis then proceeded in two integrated phases. First, a quantitative keyword analysis was conducted to objectively map the dominant thematic landscape of the regulatory discourse. Subsequently, in a qualitative phase,

[1]See Navigating the Global AI Governance Landscape – Global Partners Digital. (n.d.). Available at: https://www.gp-digital.org/navigating-the-global-ai-governance-landscape/ (accessed 18 December 2024).

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Table 1. Corpora information

Type	Corpus	Tokens
Target corpus	Corpus of International Regulations on AI	397674
Reference corpus	COCA	9412521

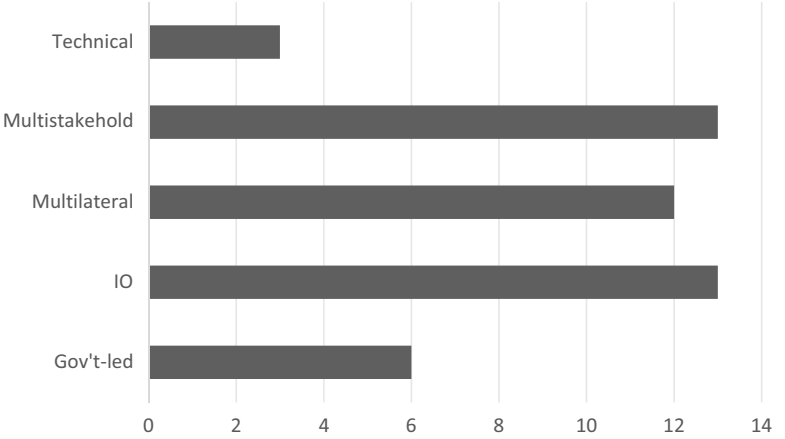


Figure 2. Distribution of drafting institutions.

these computer-identified keywords and their surrounding patterns were examined within their original textual context. This second step involved a sociosemiotic 'close reading' to interpret how key terms were used to construct meaning and frame the regulatory problem, allowing for a systematic yet nuanced interpretation of their function.

4. Results and discussion

This discussion interprets the findings of our corpus analysis through the lens of social semiotics. Viewing the collected AI normative texts as a form of discourse, we draw upon Hallidayan and Foucauldian perspectives to understand them not as objective statements, but as selective representations of reality. These texts utilise specific semiotic resources – namely, vocabulary – to construct a particular understanding of AI governance. The critical question, therefore, is how the international regulation of AI is being characterised. What specific themes and topics are brought to the forefront through the discourse's choice of language? To answer this, we turn to the keyword analysis detailed in Table 2. These keywords, generated by comparing our target corpus against the COCA reference corpus, reveal the core thematic structure of the AI regulatory discourse.

The keyword analysis reveals a distinct clustering effect, indicating that international normative documents have converged around a shared set of concepts and principles. Drawing from Van Leeuwen's social semiotic insight – that discourse selectively represents elements of a social practice – this keyword list highlights the core components prioritised within the quasi-international legal discourse of AI governance. A thematic categorisation of these keywords reveals three primary domains: (1) Core Subjects and Concepts: AI, data, system(s), technology, digital, risk(s) and development; (2) Regulatory Mechanisms and Actions: ensure, governance, framework and stakeholders; (3) Guiding Values and Principles: human, rights, principles, trustworthy, safety and innovation. Based on this categorisation, the foundational values currently driving international AI regulation are human-centricity, trustworthiness, safety and innovation.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Table 2. Keyword list of target corpus

Rank	Type	Freq_Tar	Range_Tar	NormFreq_Tar	NormRange_Tar	Keyness (Likelihood)
1	AI	8138	44	20463.998	0.936	47958.405
2	Systems	2404	41	6045.153	0.872	11592.848
3	Data	2678	43	6734.159	0.915	10514.831
4	System	1938	35	4873.338	0.745	6037.365
5	Human	1689	44	4247.197	0.936	5578.667
6	Risks	935	39	2351.172	0.83	4748.321
7	Rights	1317	40	3311.758	0.851	4576.739
8	Development	1312	46	3299.185	0.979	4223.87
9	Artificial	718	42	1805.499	0.894	3922.146
10	Governance	693	33	1742.633	0.702	3888.614
11	Principles	746	42	1875.908	0.894	3455.419
12	Risk	993	34	2497.02	0.723	3443.73
13	Use	1709	43	4297.49	0.915	3361.777
14	Intelligence	789	43	1984.037	0.915	3228.942
15	Ensure	727	44	1828.131	0.936	3133.453
16	Trustworthy	482	23	1212.048	0.489	2971.495
17	Technologies	601	40	1511.288	0.851	2836.969
18	Stakeholders	475	33	1194.446	0.702	2815.176
19	Safety	762	36	1916.142	0.766	2794.567
20	Digital	668	38	1679.768	0.809	2784.613
21	Innovation	523	29	1315.148	0.617	2631.774
22	Technology	794	39	1996.61	0.83	2421.648
23	Framework	515	34	1295.031	0.723	2407.26
24	Impact	716	37	1800.47	0.787	2392.437
25	Indicators	422	15	1061.171	0.319	2119.686

Admittedly, the analysis also identified other crucial values such as privacy, transparency, accountability, sustainability and explainability. A comprehensive ranking of these normative keywords by their statistical keyness is presented in Table 3.

The keywords related to ethical principles and values reveal a clear hierarchy within the dataset, with certain values receiving significantly more emphasis than others. This finding aligns with the evolving focus of AI principles research. For instance, a landmark qualitative study by Jobin *et al.* (2019) identified eleven ethical principles, noting an emerging global consensus around five core themes: transparency, fairness, non-maleficence, responsibility and privacy. Subsequent research by American scholars highlighted eight principles, adding safety, human control of technology, professional responsibility and human values to the list, which notably signalled a clearer and more prevalent focus on human-centric ethics (Fjeld *et al.* 2020). Similarly, Chinese scholars have summarised various frameworks, including three, four and eight-principle models that primarily feature values such as human-centricity, trustworthiness, accountability, safety, fairness and beneficence.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Table 3. Rankings of values in international AI regulations

Type	Freq_Tar	Range_Tar	NormFreq_Tar	NormRange_Tar	Keyness (Likelihood)
Human	1689	44	4247.197	0.936	5578.667
Rights	1317	40	3311.758	0.851	4576.739
Trustworthy	482	23	1212.048	0.489	2971.495
Safety	762	36	1916.142	0.766	2794.567
Innovation	523	29	1315.148	0.617	2631.774
Privacy	477	35	1199.475	0.745	2057.963
Ethical	392	36	985.732	0.766	1854.973
Transparency	310	39	779.533	0.83	1589.322
Responsible	437	39	1098.89	0.83	1374.128
Accountability	279	38	701.58	0.809	1285.985
Security	525	38	1320.177	0.809	1061.219
Discrimination	263	26	661.346	0.553	1056.98
Bias	249	34	626.141	0.723	1011.597
Explainability	154	23	387.252	0.489	987.365

Among these studies, the classification by Jobin *et al.* is particularly detailed, covering a comprehensive range of AI legal and ethical principles. Therefore, to systematically assess the hierarchy of values in the current discourse, this study adopted their framework to guide a frequency analysis of the corpus. This approach is methodologically sound, as word frequency in legislative corpora is often used to reveal ideological leanings and policy priorities (Slapin and Proksch 2014). The analysis thus provides an empirical ranking of the main principles currently emphasised in international AI governance.

A frequency analysis of the entire corpus reveals that the AI principles cluster into ten primary thematic categories (See Table 4). In descending order of frequency, these are: safety, human-centricity, fairness, beneficence, accountability, sustainability, privacy, trustworthiness, transparency and freedom. This classification largely aligns with the framework proposed by Jobin *et al.* (2019). Most notably, the terms related to the top three principles – safety, human-centric and fairness – appear with a significantly higher frequency than all others.

However, this quantitative ranking has its limits. First, such classifications are not absolute. Existing AI principle frameworks often lack systematicity and uniform expression, and different global actors promote different governance visions. This is evident in our corpus, where the 'Range' column in Table 4 shows that the distribution and density of these principle-related terms vary considerably across texts. Second, from a social semiotic perspective, even when a legal term like 'human rights' appears as the same signifier, its meaning is dynamic and context-dependent (Cheng *et al.* 2014). These two factors – inconsistent application across texts and the inherent ambiguity of legal terms – underscore the necessity of moving beyond frequency counts to analyse contextualised meaning. Therefore, the following analysis returns to the source texts to qualitatively examine the use of the three most prominent principle clusters.

4.1. Safety

'Safety' stands out as the most prominent and frequently mentioned principle in the entire corpus. To conduct an in-depth analysis of this theme, a set of nine related search terms was compiled,

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Table 4. AI principles related words identified in the target corpus

Principle	Type	Frequency_ Target	Frequency_Reference	Range_Target
Safety	safety	762	795	36
	security	525	1732	38
	safe	198	1006	27
	secure	141	304	24
	protection	402	545	33
	prevention	46	205	16
	protect	154	721	29
	prevent	95	455	26
	integrity	87	155	21
Human-centric	dignity	62	168	18
	human	1689	2310	44
Fairness	justice	133	1161	17
	fairness	169	78	27
	consistency	15	109	10
	equality	83	201	20
	equity	81	167	15
	bias	249	190	34
	discrimination	263	208	26
	diversity	154	275	28
	accessibility	69	44	14
	reversibility	3	0	1
	inclusive	201	93	28
	inclusion	92	85	25
	inclusiveness	20	5	9
	diverse	99	239	25
Beneficence	benefit	106	610	29
	innovation	523	157	29
	benefits	302	700	37
	well-being	28	7	7
Accountability	accountable	61	106	23
	responsible	437	668	39
	liability	34	97	13
	responsibility	144	606	32
	accountability	279	131	38
Sustainability	sustainable	242	140	24
	sustainability	45	38	15

(*Continued*)

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Table 4. (*Continued*)

Principle	Type	Frequency_ Target	Frequency_Reference	Range_Target
	environment	313	820	35
	energy	79	1381	20
	resources	236	837	29
Privacy	privacy	477	294	35
	personal	254	1700	25
Trustworthy	trust	186	890	28
	trustworthy	482	13	23
Transparency	transparency	310	83	39
	transparent	81	72	28
	explainability	154	0	23
	explicability	10	0	2
	understandability	5	2	4
	interpretability	33	1	9
Freedom	freedom	134	885	21
	autonomy	109	82	23
	consent	75	167	17
	empowerment	10	38	5

Table 5. Top collocations of 'protection'

Collocate	Rank	Freqency	Range	Likelihood	Effect
Data	1	271	28	778.178	3.323
Privacy	2	72	23	257.031	3.9
Personal	3	50	14	203.494	4.283
Promotion	4	24	7	145.748	5.754
Consumer	5	24	8	122.81	5.063

guided by the keywords in both our initial analysis and Jobin *et al.*'s framework: safe, safety, secure, security, protect, protection, prevent, prevention and integrity. An examination of the collocates for these terms confirmed that their semantic use throughout the corpus consistently falls within the conceptual domain of the safety principle. The emphasis on this theme is quantitatively significant. The combined frequency of these nine terms reaches 2,410 occurrences. Furthermore, their broad distribution across the documents suggests a widespread consensus on the centrality of safety among different types of actors. Specifically, the core term 'security' appears in thirty-eight of the forty-seven documents (81 per cent), while 'safety' is present in thirty-six (77 per cent).

Within the cluster of terms related to the safety principle, the pairs safety/safe and security/secure are consistently linked, necessitating a clear distinction. A contextual examination of the corpus reveals that safety typically refers to the internal integrity of an AI system and the

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

avoidance of accidental harm, whereas security pertains to its external resilience against attacks. This distinction is explicitly articulated in UNESCO's Recommendation on the Ethics of Artificial Intelligence, which states that, 'Unwanted harms (safety risks), as well as vulnerabilities to attack (security risks) should be avoided . . . '. This conceptual difference is also reflected in their distribution and collocates. Safety appears most frequently in technical, standard-setting documents and its top collocate is function, reinforcing its connection to internal operations. In contrast, security is distributed broadly across all document types, and its most significant collocate is national, pointing to a different set of primary concerns. The prominence of national security suggests that the discourse on AI safety is deeply embedded in a framework of legally protected interests. This securitisation of AI can be understood as part of a broader logic of governance for managing uncertain futures, where principles of precaution, pre-emption and preparedness become central. The focus on 'safety' and 'risk' in the texts reflects a political desire to anticipate and neutralise potential threats before they materialise, a key feature of contemporary governance imaginaries (Lazaro and Rizzi 2023). This framework can be conceptualised as a three-tiered model of 'National Security – Social Security – Individual Security', an evolution from the traditional Roman law dichotomy of public and private law. In this model, national security is the broadest level, concerning state sovereignty. Social security is the intermediate level, relating to public order and the stability of societal domains. Individual security is the narrowest, encompassing not only traditional personal and property safety but also emerging technological and industrial safety. The development and use of AI implicate all three levels of this security framework, from public interests to private ones. Because the authors in our corpus – ranging from inter-governmental organisations and states to NGOs and private firms – each selectively emphasise different facets of security, the resulting discourse is inherently multi-faceted, regionalised and context-dependent. Consequently, any effort to build a meaningful global consensus must be founded upon a clear acknowledgement of these diverse interests to fully define the principle's intension and extension. In this context, the 'safety' discourse within AI soft law documents can be understood as what the legal sociologist Philip Selznick termed 'incipient law' – patterns and claims 'latent in the evolving social and economic order' that are highly probable to receive legal recognition in due course. The widespread consensus on safety, emerging from the practices and declarations of both state and non-state actors, forms a normative foundation from which future binding treaties or customary international law on AI safety may evolve. The process of securitisation, therefore, is not just a political manoeuver; it can be seen as a jurisgenerative process where social anxieties about risk are potentially translated into the building blocks of a new legal order.

'Protection' is another key term within the safety principle, revealing the primary objects of this concern. The collocation results for protection (Table 5) are dominated by data, privacy, personal and consumer, outlining a clear focus on safeguarding the data and privacy of individuals. This focus is consistently reinforced in the source texts, such as the G20's call to 'promote the protection of privacy and personal data' and UNESCO's mandate that data processing must align with 'data protection and data privacy' laws and principles. Beyond what is protected, the discourse reveals a crucial insight into how protection is framed. Drawing on Foucault's (2013) concept of discourse as a socially constructed knowledge and Van Leeuwen's (2005) notion of 'substitution', we can observe a key discursive strategy at play: nominalisation. Van Leeuwen argues that replacing verbs (doing) with nouns (being) objectifies an issue, framing it as a settled state. This is precisely what occurs here: the nominalised forms protection and prevention appear with far greater frequency than their verb counterparts protect and prevent. This suggests the discourse treats data and privacy protection not as an ongoing action to be advocated for, but as an already established consensus – a static feature of the governance landscape. This discursive emphasis on protection, however, exists in tension with the practical reality of AI development, where the collection and use of data serve as the technology's essential fuel. While data, algorithms and computing power are all core elements of AI, massive datasets are what enabled the recent

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

'emergent intelligence' in models like ChatGPT, whose core algorithms have existed for years. A final analysis of the word data itself within our corpus confirms the discursive priority: protection co-occurs with data far more frequently than terms related to data utilisation, such as collection, training, sharing or processing. This suggests a central finding: while data use is paramount for technological advancement, data protection appears to constitute the core theme of the regulatory discourse.

The pronounced emphasis on 'safety' is one of this study's most striking findings. In contrast to earlier analyses of global AI governance datasets, the safety principle is now discussed more frequently, suggesting a new global consensus on its importance. This evolution can be understood through Balkin's (1990) concept of 'ideological drift'. Balkin argues that as political and legal concepts are expressed through signs, their ideological leanings can shift over time as those signs are iterated in new contexts. The rising prominence of 'safety' thus demonstrates this diachronic function of signs, where the discourse around AI principles has evolved in just a few years.

Furthermore, the critical nature of social semiotics can reveal previously unrecognised conflicts and prompt a reassessment of existing intuitions. Paul's (2023) semiotic study of the EU AI Act provides a powerful recent example. Paul deconstructs the EU's 'risk-based' approach, arguing – based on legislative analysis and expert interviews – that the risk classifications were not objective but were rather a 'political decision and value statement'. This framing was a strategic discursive choice that allowed the EU's regulatory framework to 'stand out' in the global competition over AI norms. Similarly, the discursive construction of the 'safety' principle in our corpus merits critical reflection. A sociosemiotic perspective prompts further questions: how, precisely, was the regulation of AI 'securitized?' Whose viewpoints does this safety-centric discourse represent and amplify? And through what semiotic processes was this principle generated and propagated?

While fully answering these questions requires more contextual material, the current international political climate offers potential explanations. One is the broader trend of 'pan-securitization', where numerous international issues are framed through a security lens. For a disruptive technology like AI, a 'safety narrative' offers international actors more control than a 'development narrative'. Second, the still-maturing international AI regime is heavily influenced by the domestic laws of first-mover jurisdictions, and these legal systems have themselves been observed to be undergoing a clear securitisation process (Mügge 2023; Zeng 2021).

4.2. Human-centric

Alongside safety, a human-centric approach emerges as one of the most significant principle themes in the corpus. This is immediately evident from the prominence of the keyword human, which, with a frequency of 1,686, is the most common of all principle-related terms. Its statistical keyness is second only to keywords that refer to AI itself. A qualitative review of the corpus shows that human combines with other words to form a rich semantic network encompassing concepts like rights, dignity and well-being. Therefore, this analysis uses the term 'human-centric' – a label common in both international law and contemporary AI ethics scholarship (Amariles and Baquero 2023; Rožanec *et al.* 2023) – as an umbrella concept to encapsulate the diverse principles revolving around the keyword human. To deconstruct this concept, a collocate analysis of human was performed, with the results presented in Table 6. Additionally, the Cluster function was used to identify common two-word phrases (n-grams) involving human, as shown in Table 7.

The analysis of human's collocates in Table 6 and word clusters in Table 7 reveals that the 'human-centric' principle encompasses two primary dimensions. The first dimension concerns the relationship between humans and AI systems. A contextual examination of key collocates like loop, oversight and autonomy confirms that their semantic use consistently points to human interaction with and control over AI. This theme is further reinforced by the prominence of two-

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Table 6. Top collocations of 'human'

Collocate	Rank	Frequency	Range	Likelihood	Effect
Rights	1	603	35	1814.592	3.43
Loop	3	72	6	355.964	4.869
Intelligent	4	95	2	333.632	3.833
Well	5	217	15	315.673	2.154
Oversight	6	100	20	303.101	3.466
Dignity	7	65	13	298.823	4.626
Autonomous	8	95	5	272.363	3.339
Autonomy	10	60	12	199.744	3.696
Human	11	206	20	168.241	1.522
Freedoms	12	46	10	159.162	3.796

Table 7. Top clusters of 'human'

Cluster	Rank	Freq	Range
Human rights	1	531	33
Human well	4	173	4
Human oversight	10	51	18
International human	11	43	15
Human dignity	12	39	11
Human beings	13	35	10
Human centric	14	34	10
Human autonomy	16	33	7
Human agency	18	30	9
Human intervention	19	24	11

word clusters such as human agency and human intervention. The second dimension is an emphasis on human values, highlighted by collocates like dignity, freedoms and well-being. Within this dimension, the most significant finding is the deep integration of the international human rights framework. The high co-occurrence score of the collocate rights, combined with the statistical prominence of the cluster human rights, demonstrates that international human rights law has become a foundational component of the current AI regulatory discourse.

Regarding the human–AI relationship, the corpus reveals a strong emphasis on human control over technology. This control manifests as: human oversight of automated decisions; human independence within the decision-making loop (ensuring autonomy and agency); and the capacity for direct human intervention. The concept of 'oversight' primarily refers to the human review of an AI system's decisions, typically conducted *ex post* to correct outcomes. This requires that humans retain effective autonomy, as noted in the *Principles for Responsible AI Innovation*: 'human control and oversight are the ability and opportunity for humans to adequately supervise, engage and interfere with an AI system . . . law enforcement agencies are advised to verify that . . . humans have the last word'. Documents like the *Blueprint for an AI Bill of Rights* further specify that such oversight should be enhanced in sensitive domains like the criminal justice system,

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

employment and health care. However, despite this emphasis, the requirements for human oversight are often described in general terms, lacking detailed scenarios or clear hierarchical rules. This is reflected in practice, where legal professionals acknowledge the necessity of verifying AI-generated outputs to mitigate risks like 'hallucinations', but formal training on how to do so remains scarce. The reliance on human oversight, therefore, becomes a matter of individual professional responsibility rather than a systematically implemented rule (Terzidou 2025). Different documents propose varying levels of review, typically based on risk. For example, the EU's *Ethics Guidelines for Trustworthy AI* notes that oversight can be required 'in varying degrees . . . depending on the . . . application area and potential risk', and operationalises this through a questionnaire-style self-assessment list. While flexible, this approach delegates the interpretation of oversight intensity and mechanisms to the implementing entity. This may lead to a diversification of standards in practice, undermining the goal of achieving international consensus. Such vagueness is characteristic of the current international AI regime, which is dominated by non-binding soft law instruments (Elias and Lim 1997). This implementation challenge is rooted in the nature of international legal communication. As a cultural semiotic system, international law must be interpreted by a diverse global community (the 'receivers') who may not share the same context as the drafters (the 'senders'). The comparability and effectiveness of any legal regime depend on a shared interpretation of meaning (Anne *et al.* 2020). The inherent ambiguity of soft law, while inclusive, exacerbates this problem, as it can lead to divergent practices that reduce comparability. Therefore, from the perspective of ensuring clear semiotic communication and robust dialogue, hard law mechanisms – such as the Council of Europe's *Framework Convention on Artificial Intelligence* – represent a crucial path forward for achieving a more thoroughly negotiated and stable consensus.

In the human–AI relationship, the concepts of human independence, oversight and control are tightly interwoven. The keyword loop, for instance, is central to a tripartite classification found in roughly a quarter of the documents: human-in-the-loop, human-over-the-loop and human-out-of-the-loop, which represent escalating levels of automation and risk.[2] A core tenet of human independence is that even in human-out-of-the-loop scenarios, a person retains the ability to exit the automated process. In the corpus, this capacity is represented by human autonomy and human agency. The *Ethics Guidelines for Trustworthy AI* define human agency as the ability for users to 'make informed autonomous decisions regarding AI systems', while UNESCO's *Recommendation on the Ethics of Artificial Intelligence* links human autonomy and human agency directly to the protection of human dignity. However, the principles of control and intervention reveal a significant tension at the heart of AI governance. While regulatory texts call for human control, technical standards often define autonomous systems in contradictory terms. For example, the *IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being* defines an intelligent system as one that can 'carry out some task with or without limited human intervention capable of decision making by independent inference'. This characterisation of AI as inherently free from intervention directly conflicts with the human-centric demand for control. This contradiction likely stems from what appears to be one of the deepest conflicts within AI regulation: the tension between development and safety. This development-versus-safety conflict brings the second dimension of the human-centric principle – human values – into sharp focus. On the other side of the risks posed by AI is its immense potential to benefit society; collocates of human show that 'development' and 'well-being' are also

[2]According to the ASEAN Guide on AI Governance and Ethics, human-in-the-loop means the AI system only provides recommendations while the human retains full control over the decision. In a human-over-the-loop model, a person acts as a supervisor who monitors the AI and can intervene if its actions are undesirable, similar to how a driver must remain alert in a self-driving car to take over when needed. Conversely, human-out-of-the-loop refers to an AI system that executes decisions with full autonomy and without any human selection or approval, such as a recommendation algorithm that pushes content to users on its own.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

treated as key legal interests. The governance challenge, therefore, is a fundamental value conflict between the legal interest of safety and that of development. As long as different legislative contexts prioritise these competing values differently, the formation of an effective and stable global consensus will remain a significant challenge.

One of the most significant findings relating to the human-centric principle is the widespread integration of the international human rights law (IHRL) framework, mentioned in thirty-three of the forty-seven documents. This trend is not isolated to AI governance; it is part of a broader 'transnational human rights' movement where international norms, constitutional rights and humanitarian law increasingly interact and shape both state and non-state action (Klug 2005). This is supported by the high statistical keyness of the terms human and rights and the frequent human rights law cluster. The contextual analysis shows that human rights has a dual meaning: it refers both to core values, often listed with dignity and fundamental freedoms, and to the formal legal framework itself. These two meanings are tightly linked; when documents elaborate on values, they consistently align them with IHRL. For instance, the Principles for Responsible AI Innovation states that its concepts 'are derived from the Universal Declaration of Human Rights', while other reports use specific ICCPR articles to define principles like non-discrimination, directly aligning AI ethics with established human rights law. The broad application of IHRL signals a growing recognition of its utility for AI governance. The IHRL framework provides a mature, systematic and universal methodology for assessing and regulating AI, built on decades of global consensus and already integrated into national legal systems (Rodrigues 2020). Its suitability stems from its focus on human dignity and its clear delineation of state and private sector responsibilities (Donahoe and Metzger 2019). The direct applicability of this framework is clear when mapping cornerstone documents like the *Universal Declaration of Human Rights* (UDHR) to key AI principles. For example, the UDHR's principles of equality and non-discrimination (Article 2) directly address AI fairness; the rights to remedy and due process (Articles 8–11) provide a foundation for transparency and accountability; and the right to privacy (Article 12) anchors data protection requirements. Ultimately, the relationship between the nascent discourse of AI regulation and the established IHRL framework can be understood as a form of 'manifest intertextuality' (Fairclough 1992). The two function as interconnected legal semiotic systems at the international level. The semiotic meaning of AI regulation is endowed and specified through its dialogue with IHRL. Concurrently, the semiotic meaning of the human rights framework is itself reinterpreted and enriched as it is applied to the new and complex problem domain of artificial intelligence, suggesting a dynamic and mutually constitutive relationship.

4.3. Fairness

As one of the three most significant principles revealed in the corpus, fairness is distinguished by having the richest set of associated keywords, indicating its broad semantic range. The centrality of this theme is underscored by the fact that its antithesis, discrimination, is mentioned in thirty-four of the forty-seven documents. The keywords encompassed by this principle include fairness, justice, equality, diversity, inclusion, consistency, accessibility, unbiased and non-discrimination. As a deontological principle, fairness is an essential requirement for 'AI for good'. However, it is also notoriously vague, containing multiple, often competing, meanings that can lead to value conflicts. This presents a challenge for purely qualitative analysis, as noted in the Artificial Intelligence Risk Management Framework: 'Standards of fairness can be complex and difficult to define because perceptions of fairness differ among cultures and may shift depending on application'. To navigate this complexity, this study utilises corpus methods (frequency and keyness calculations) to distil the 'greatest common denominator' from the discourse across the extensive document set. Based on semantic proximity, keywords under principles of fairness in Table 4 can be grouped into four primary thematic categories: (1) Non-discrimination and Non-bias (non-discrimination, unbiased); (2) Fairness and Justice (fairness, justice, equity); (3) Equality

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

and Accessibility (equality, accessibility, with equity also being relevant); and (4) Diversity and Inclusion (diversity, inclusion). Based on a ranking of their combined frequency and keyness, 'Non-discrimination and Non-bias' emerges as the most significant dimension of the fairness principle. This is followed, in order, by 'Fairness and Justice', 'Diversity and Inclusion' and finally 'Equality and Accessibility'. The following analysis will further examine these quantitative discourse results.

Non-discrimination and the prevention of bias represent the most significant dimension of the fairness principle, with discussions of bias found in thirty-four of the forty-seven documents and non-discrimination in twenty-six. The discourse is primarily concerned with algorithmic bias, which is defined as 'systematic and repeatable errors in a computer system that create unfair outcomes' in the *Assessment List for Trustworthy AI*. The root causes are identified as flawed or unrepresentative training data and poor algorithmic design. As many documents note, the harms of such bias are amplified in sensitive domains, with frameworks like the *AI Safety Governance Framework* stating that 'poor-quality datasets can lead to biased or discriminatory outcomes . . . regarding ethnicity, religion, nationality and region'. However, a key limitation of the current discourse is that it concentrates on bias as a problem to be avoided rather than articulating a clear, operational vision for an unbiased AI system. This stems from the semantic uncertainty of the term bias itself. Most texts address discrimination in general terms, often using vague modifiers like harmful, unjust or unintended. Even the most precise definitions, such as that in the EU's *Ethics Guidelines*, offer only a broad framing of bias as 'an inclination of prejudice towards or against a person, object, or position', without providing concrete operational criteria. This lack of specificity can be explained by the inherent distinction between the legal and moral dimensions of discrimination. The moral understanding of bias is often unstable and subject to rapid semantic shifts based on social trends. Law, adhering to a principle of contraction, is inherently cautious about codifying such fluid moral concepts. This challenge, combined with the fact that the meaning of bias is spatiotemporally dependent – varying significantly across different regions and legal cultures – results in a principle that is widely endorsed in theory yet lacks the concrete substance required for effective legal practice.

The collocation analysis shows that fairness and justice are core components of the broader fairness principle, frequently appearing alongside non-discrimination. Similar to the other concepts, it has multiple meanings. Documents like *A pro-innovation approach to AI regulation* acknowledge this complexity, suggesting that fairness should be interpreted in line with existing legal frameworks such as human rights law. A particularly useful approach, proposed in the EU's *Ethics Guidelines for Trustworthy AI* and supported by a contextual review of the corpus, is to divide the concept into substantive fairness and procedural fairness. Substantive fairness, though often defined in general terms, appears widely. For example, the *Principles for Responsible AI Innovation* defines it as ensuring 'a just and non-discriminatory treatment of individuals and groups and a contribution to a more equitable society'. Procedural fairness complements this by safeguarding individual recourse. The same documents emphasise that procedural fairness 'requires that agencies safeguard people's ability to contest decisions supported by AI systems and to be compensated if such decisions are harmful to them'. This distinction highlights a key dynamic: because the concept of fairness and justice in legislative documents remains generalised and slow to evolve, enforcement and judicial action often take the lead in guaranteeing fairness in the AI domain. However, a significant gap remains, as these enforcement-led approaches currently lack effective, detailed guidance on how to reconcile the competing values of fairness and innovation.

Inclusion, primarily represented by the keywords diversity and inclusion, which often appear together, is another important dimension of the fairness principle. A contextual analysis reveals that inclusion is applied in two primary scenarios. The first is inclusion in the design of AI systems, which requires that development processes be multi-faceted and participatory. As the EU's Ethics Guidelines for Trustworthy AI states, 'It is critical that . . . the teams that design,

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

develop, test and maintain . . . these systems reflect the diversity of users and of society in general', including different genders, races, nationalities and cultures. This extends beyond the design teams to also consider the diversity of the technology's intended audience. Furthermore, organisations are called upon to support this participation, with the *Partnership on AI* noting that 'committing organizational resources demonstrates a different level of dedication to pursuing inclusively . . . developed technology'. The second scenario is the inclusive sharing of AI's benefits. This interpretation of inclusion is widely mentioned throughout the corpus and serves to link the principle of fairness with the broader goals of human-centricity and trustworthy AI. For example, the OECD's *Policy and Investment Recommendations for Trustworthy AI* explains inclusion as ensuring that all people can benefit from artificial intelligence and that it contributes to social well-being.

Finally, the cluster of words represented by 'equality' has a certain semantic similarity to 'fairness and justice', but from the context of the signs, equality has a different use case compared to 'fairness'. 'Equality' is more absolute, meaning that regardless of whether the situations of different groups are similar, they should receive the same opportunities and protection, whereas fairness is more about adapting to local conditions, focusing on guaranteeing the treatment and opportunities of similar groups. In the corpus, the most common collocates of 'equality' include 'race', 'gender', 'dignity', 'economic' and 'rights'. Taking 'race' and 'gender' as examples, even if there are large differences between individuals, they are 'absolute' equality matters in the AI context. In addition, the principle of 'equality' is often mentioned together with the principle of 'non-discrimination', but at the semantic level, the principle of 'equality' goes a step further than the principle of 'non-discrimination'. In the AI context, non-discrimination still retains the possibility of treating different situations differently based on objective reasons, whereas equality means applying the same rules – for example, the same algorithm rules – so that all groups have equal rights to information, data, knowledge, market access and the distribution of utility (Yu 2020). In the corpus documents, equality in the AI context mainly includes three aspects: equality from a human rights perspective (race, gender), equality in the accessibility of technology and equality in sharing the results of AI technology. For example, the Recommendation on the *Ethics of Artificial Intelligence* released by UNESCO mentions that the goals of the recommendation include equality from a human rights perspective, 'to protect, promote and respect human rights and fundamental freedoms, human dignity and equality, including gender equality', as well as equality in accessing AI knowledge and in sharing the results of AI development, 'to promote equitable access to developments and knowledge in the field of AI and the sharing of benefits'.

5. Conclusion

From a social semiotic perspective, this paper presents a systematic analysis of consensus-building and legitimisation within the global governance of AI. Based on a corpus of forty-seven international normative documents, this study identifies three primary characteristics of the current landscape: a fragmented regulatory architecture, a normative system dominated by soft law and the dynamic participation of multiple actors. The analysis reveals that the three core principles of 'safety', 'human-centricity' and 'fairness' form a cross-textual consensus through semiotic expression. 'Safety' constructs a hierarchical framework of legally protected interests, from the national to the individual level; 'human-centricity' emphasises technological controllability through signs like 'human rights' and 'oversight'; and 'fairness' seeks to balance social justice through discourses of 'non-discrimination' and 'diversity'.

However, the pervasive use of nominalisation in these soft law texts (e.g. 'protection' instead of 'to protect'), while effective at building surface-level consensus, also exposes a core tension between this emerging 'living law' and the state's formal 'norms for decision'. This ambiguity undermines practical enforcement, leading to fragmented standards and interpretive disagreements during implementation. The social semiotic framework further demonstrates that these

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

governance texts are not merely technical products but are also mappings of global power relations. For instance, the prevalence of 'securitization' discourse reflects the risk-control imperatives of sovereign states in technological competition, a logic of governance increasingly shaped by a desire to pre-empt uncertain futures. At the same time, the frequent citation of the 'human rights framework' highlights a strategy by non-state actors to contest discursive power using established legal signs, a process central to the formation of transnational legal fields.

At a theoretical level, this study offers a certain contribution by moving beyond the traditional hard law–soft law dichotomy. By treating AI governance texts as a dynamic semiotic system, it demonstrates how a new transnational legal order for AI is being constructed, one characterised by the features of legal pluralism. This approach aligns with socio-legal scholarship that emphasises 'living law' over state-centric models and aims to contribute an analytical tool for understanding the legitimisation of technology governance. The challenge for the future of international law, therefore, is whether and how this vibrant, pluralistic 'living law' of AI governance will be integrated and made accountable through more robust legal mechanisms, without stifling the innovation and multi-stakeholder collaboration that gave it rise.

It must be acknowledged that this study is limited by the timeliness of its corpus and sample size, and has not fully captured the impact of rapid technological iteration on governance discourse. Furthermore, an analysis of non-textual semiotic practices awaits future research. Such research could be expanded to include multimodal signs (e.g. images, algorithmic code) or to track the chronological evolution of these norms, offering a more comprehensive response to the globality and complexity of AI governance. In addition, as global AI legislation evolves and more comparable soft law and hard law corpora become available, this move beyond the dichotomy can evolve into comparative research within these respective datasets, seeking to bridge and reconcile the potential value differences that may exist between them.

Overall, the global regulation of AI is not merely a technical challenge but a profound semiotic and legal process laden with political and moral disagreements. A durable balance between common human values and technological transformation can only be achieved by embracing difference within consensus and upholding fundamental principles amid innovation, recognising the complex interplay between the living law of society and the formal law of states.

Acknowledgements. This work was supported by the National Social Science Fund of China (No. 24BYY151) under the project "Corpus-based Research on Cyber Security Discourse System". We would like to express our sincere gratitude to the anonymous reviewers for their inspiring feedbacks and encouraging comments.

Competing interests. No competing interests.

References

Amariles DR and Baquero PM (2023) Promises and limits of law for a human-centric artificial intelligence. *Computer Law & Security Review* **48**, 105795.

Anne W, Matulewska A and Cheng L (2020) Law as a culturally constituted sign-system – A space for interpretation. *International Journal of Legal Discourse* **5**(2), 239–267.

Balkin JM (1990) The promise of legal semiotics. *Texas Law Review* **69**, 1831.

Berger P and Luckmann T (2016) The social construction of reality. In Longhofer W and Winchester D (eds.), *Social Theory Re-Wired*. London: Routledge, pp. 92–101.

Burley AMS (1993) International law and international relations theory: A dual agenda. *American Journal of International Law* **87**(2), 205–239.

Cheng L, Cheng W and Sin KK (2014) Revisiting legal terms: A semiotic perspective. *Semiotica* **2014**(202), 167–182.

Cheng L and Danesi M (2019) Exploring legal discourse: a sociosemiotic (re)construction. *Social Semiotics* **29**(3), 279–285.

Cheng L and Sin KK (2008) Terminological equivalence in legal translation: A semiotic approach. *Semiotica* **2008**(172), 33–45.

Cheng L, Sun Y and Chen F (2019) Cultural heritage in legal settings: A semiotic analysis. *Zeitschrift für Semiotik* **41**(3–4), 131–151.

Clarke R (2019) Regulatory alternatives for AI. *Computer Law & Security Review* **35**(4), 398–409.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Cobley P (2010) *The Routledge Companion to Semiotics*. London: Routledge.
Cobley P and Randviir A (2009) Introduction: What is sociosemiotics?. *Semiotica* **2009**(173), 1–39.
De Saussure F (2004) Course in general linguistics. In Rivkin J and Ryan M (eds.), *Literary Theory: An anthology*. Oxford: Blackwell Publishing, pp. 59–71.
Donahoe E and Metzger MM (2019) Artificial Intelligence and Human Rights. *Journal of Democracy* **30**(2), 115–126.
Dwivedi YK, Hughes L, Ismagilova E, Aarts G, Coombs C, Crick T, Duan Y, Dwivedi R, Edwards J, Eirug A, Galanos V, Ilavarasan PV, Janssen M, Jones P, Kar AK, Kizgin H, Kronemann B, Lal B, Lucini B, Medaglia R, Le Meunier-FitzHugh K, Le Meunier-FitzHugh LC, Misra S, Mogaji E, Sharma SK, Singh JB, Raghavan V, Raman R, Rana NP, Samothrakis S, Spencer J, Tamilmani K, Tubadji A, Walton P and Williams MD (2021) Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* **57**, 101994.
Ehrlich E and Ziegert KA (2017) *Fundamental Principles of the Sociology of Law*. New York: Routledge.
Elias O and Lim C (1997) 'General principles of law', 'soft' law and the identification of international law. *Netherlands Yearbook of International Law* **28**, 3–49.
Erman E and Furendal M (2022) The global governance of artificial intelligence: Some normative concerns. *Moral Philosophy and Politics* **9**(2), 267–291.
Evans MD (2014) *International Law*. Oxford: Oxford University Press.
Fairclough N (1992) *Discourse and social change*. Cambridge: Polity Press.
Fairclough N (2003) *Analysing discourse*. London: Routledge.
Fjeld J, Achten N, Hilligoss H, Nagy A and Srikumar M (2020) Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center Research Publication (2020–2021).
Foucault M (2013) *Archaeology of Knowledge*. London: Routledge.
Fui-Hoon Nah F, Zheng R, Cai J, Siau K and Chen L (2023) Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research* **25**(3), 277–304.
Goździ-Roszkowski S (2021) Corpus Linguistics in Legal Discourse. *International Journal for the Semiotics of Law - Revue internationale de Sémiotique juridique* **34**(5), 1515–1540.
Green JF and Auld G (2017) Unbundling the regime complex: The effects of private authority. *Transnational Environmental Law* **6**(2), 259–284.
Günther K (2008) Legal pluralism or uniform concept of law? Available at: https://scholar.archive.org/work/u7rpkfj2tbg4rdfoyez4sq5vsm/access/wayback/http://www.nofoundations.com/issues/NoFo5Gunther.pdf.
Haggard S and Simmons BA (1987) Theories of international regimes. *International Organization* **41**(3), 491–517.
Halliday MAK (2014) *Language as Social Semiotic: The Discourse Studies Reader*. Amsterdam: John Benjamins, pp. 263–272.
Hertogh M (2008a) *Living Law: Reconsidering Eugen Ehrlich*. London: Bloomsbury Publishing.
Hertogh M (2008b) What is non-state law? Mapping the other hemisphere of the legal world. In Van Schooten H and Verschuuren J (eds.), *International Governance and Law: State Regulation and Non-state Law*. Cheltenham: Cheltenham, pp. 11–30.
Hodge R and Kress G (1988) *Social Semiotics*. Cambridge: Polity Press.
Hymes D (1971) The contribution of folklore to sociolinguistic research. *The Journal of American Folklore* **84**, 42–50.
Igbinenikaro E and Adewusi AO (2024) Navigating the legal complexities of artificial intelligence in global trade agreements. *International Journal of Applied Research in Social Sciences* **6**(4), 488–505.
Jackson BS (1985) Semiotics and legal theory. Available at: https://philpapers.org/rec/JACSAL.
Jobin A, Ienca M and Vayena E (2019) Artificial intelligence: the global landscape of ethics guidelines. *Nature Machine Intelligence* **1**(9), 389–399.
Jonas T, Eva E, Markus F, Johannes G, Mark K and Magnus L (2023) The global governance of artificial intelligence: Next steps for empirical and normative research. *International Studies Review* **25**(3), viad040.
Kennedy D (1991) A semiotics of legal argument. *Syracuse Law Review* **42**, 75.
Kennedy DW (1980) Theses about international law discourse. Available at: https://dash.harvard.edu/bitstream/handle/1/30205767/ThesisInterntnlLaw1980_20100618165429.pdf?sequence = 1.
Kevelson R (1988) *The Law as a System of Signs*. Boston, MA: Springer.
Klabbers (1998) The undesirability of soft law. *Nordic Journal of International Law* **67**(4), 381–391.
Klug H (2005) Transnational human rights: Exploring the Persistence and globalization of human rights. *Annual Review of Law and Social Science* **1**, 85–103.
Koskenniemi M (2006) *From Apology to Utopia: The Structure of International Legal Argument*. Cambridge: Cambridge University Press.
Krasner SD (1982) Structural causes and regime consequences: Regimes as intervening variables. *International organization* **36**(2), 185–205.
Kratochwil FV (1991) *Rules, Norms, and Decisions: On the Conditions of Practical and Legal Reasoning in International Relations and Domestic Affairs*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Lazaro C and Rizzi M (2023) Predictive analytics and governance: a new sociotechnical imaginary for uncertain futures. *International Journal of Law in Context* **19**(1), 70–90.

Levi-Faur D (2012) From "big government" to "big governance"? In Levi-Faur D (ed.), *The Oxford Handbook of Governance*. Oxford: Oxford University Press, pp. 3–18.

Luhmann N (1995) *Social Systems*. California: Stanford University Press.

Marchant GE and Allenby B (2017) Soft law: New tools for governing emerging technologies. *Bulletin of the Atomic Scientists* **73**(2), 108–114.

McGregor L, Murray D and Ng V (2019) International human rights law as a framework for algorithmic accountability. *International & Comparative Law Quarterly* **68**(2), 309–343.

McGrew A (2004) Power shift: from national government to global governance? In David H (ed.), *A Globalizing World?* London: Routledge, pp. 123–159.

Merry SE (1988) Legal pluralism. *Law & Society Review* **22**(5), 869–896.

Mügge D (2023) The securitization of the EU's digital tech regulation. *Journal of European Public Policy* **30**(7), 1431–1446.

Nelken D (2008) Eugen Ehrlich, living law, and plural legalities. *Theoretical Inquiries in Law* **9**(2), 443–471.

Onuf N (1985) Do rules say what they do-from ordinary language to international law. *Harvard International Law Journal* **26**, 385.

Paul R (2023) European artificial intelligence "trusted throughout the world": Risk-based regulation and the fashioning of a competitive common AI market. *Regulation & Governance* **18**(4), 1065–1082.

Rajkovic NM (2025) What is "the global"?: Reassembling how international lawyers see space and time. *International Journal of Law in Context*, 1–10.

Rodrigues R (2020) Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology* **4**, 100005.

Rožanec JM, Novalija I, Zajec P, Kenda K, Tavakoli Ghinani H, Suh S, Veliou E, Papamartzivanos D, Giannetsos T, Menesidou SA, Alonso R, Cauli N, Meloni A, Recupero DR, Kyriazis D, Sofianidis G, Theodoropoulos S, Fortuna B, Mladenić D and Soldatos J (2023) Human-centric artificial intelligence architecture for industry 5.0 applications. *International Journal of Production Research* **61**(20), 6847–6872.

Schiff D, Biddle J, Borenstein J and Laas K (2020) What's next for AI ethics, policy, and governance? A global overview. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, pp. 153–158.

Schmitt L (2022) Mapping global AI governance: A nascent regime in a fragmented landscape. *AI and Ethics* **2**(2), 303–314.

Slapin JB and Proksch SO (2014) Words as data: Content analysis in legislative studies. In Martin S, Saalfeld T and Strøm KW (eds.), *The Oxford Handbook of Legislative Studies*. Oxford: Oxford University Press, pp. 126–144

Smuha NA (2021) From a 'race to AI' to a 'race to AI regulation': Regulatory competition for artificial intelligence, law. *Innovation and Technology* **13**(1), 57–84.

Solan LM and Gales T (2017) Corpus linguistics as a tool in legal interpretation. *BYU Law Review*, 1311.

Taylor CR (2019) Editorial: Artificial intelligence, customized communications, privacy, and the General Data Protection Regulation (GDPR). *International Journal of Advertising* **38**(5), 649–650.

Terzidou K (2025) Generative AI systems in legal practice offering quality legal services while upholding legal ethics. *International Journal of Law in Context*, 1–22.

Teubner G (1991) The two faces of Janus: rethinking legal pluralism. *Cardozo Law Review* **13**, 1443.

Vaara E, Tienari J and Laurila J (2006) Pulp and paper fiction: On the discursive legitimation of global industrial restructuring. *Organization Studies* **27**(6), 789–813.

Van Leeuwen T (2005) *Introducing Social Semiotics*. London: Routledge.

Van Leeuwen T and Wodak R (1999) Legitimizing immigration control: A discourse-historical analysis. *Discourse Studies* **1**(1), 83–118.

Veale M, Matus K and Gorwa R (2023) AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science* **19**, 255–275.

Villarino JMB y (2023) Global standard-setting for artificial intelligence: Para-regulating international law for AI? *The Australian Year Book of International Law Online* **41**(1), 157–181.

Vogel F, Hamann H and Gauer I (2018) Computer-assisted legal linguistics: Corpus analysis as a new tool for legal studies. *Law & Social Inquiry* **43**(4), 1340–1363.

Weil P (1983) Towards relative normativity in international law? *American Journal of International Law* **77**(3), 413–442.

Wirtz BW, Weyerer JC and Geyer C (2019) Artificial intelligence and the public sector—Applications and challenges. *International Journal of Public Administration* **42**(7), 596–615.

Yu PK (2020) The algorithmic divide and equality in the age of artificial intelligence. *Florida Law Review* **72**, 331.

Zeng J (2021) Securitization of artificial intelligence in China. *The Chinese Journal of International Politics* **14**(3), 417–445.

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

Appendix A: List of international regulations on AI governance

Document name	Forum/process	Institutional format
Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law	Council of Europe	Multilateral
G7 Leaders' Statement on the Hiroshima AI Process	Group of Seven	Multilateral
Interim Report: Governing AI for Humanity	High-Level Multi-stakeholder Advisory Body on AI	IO
Recommendation on the Ethics of Artificial Intelligence	UN Educational, Scientific and Cultural Organization	IO
Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment	UN Human Rights Council	IO
TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management	EU-U.S. Tech & Trade Council	Gov't-led
Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence	U.S. White House/Leading AI Companies	Multi-stakehold
BLUEPRINT FOR AN AI BILL OF RIGHTS		
Artificial Intelligence Risk Management Framework		
Draft Guidance document: Ethics and governance of artificial intelligence for health	World Health Organization	IO
Global Digital Compact	UN Roadmap for Digital Cooperation/Global Digital Compact	IO
A Framework for the International Governance of AI	Carnegie Council for Ethics in International Affairs	Multi-stakehold
FOC Joint Statement on Artificial Intelligence and Human Rights	Freedom Online Coalition	Multilateral
Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects	High Contracting Parties to the Convention on Certain Conventional	IO
The Bletchley Declaration	UK Global AI Safety Summit	Gov't-led
A pro-innovation approach to AI regulation	UK Government	Gov't-led
G20 Ministerial Statement on Trade and Digital Economy	The Group of Twenty	Multilateral
Universal Guidelines for Artificial Intelligence		
2023 Ministerial Declaration	Global Partnership on AI	Multi-stakehold
Scaling Responsible AI Solutions Challenges and Opportunities		
Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain	Europol	Multilateral
Making AI Inclusive Four Guiding Principles for Ethical Engagement	Partnership on AI	Multi-stakehold
ISO/IEC TR 5469:2024	ISO/IEC	Technical

(*Continued*)

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press

(*Continued*)

Document name	Forum/process	Institutional format
ISO/IEC 23894:2023		
IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being	Institute of Electrical and Electronics Engineers Standards Association	Technical
Recommendation of the Council on OECD Legal Instruments Artificial Intelligence	OECD	Multilateral
REAIM Call to Action	Government of Netherlands	Gov't-led
Principles on Artificial Intelligence for Biodiversity Conservation	AI for Sustainable Development Goals Think Tank	Multi-stakehold
Principles on Symbiosis for Natural Life and Living AI		
Affirmative Action for Algorithms Artificial intelligence, Machine Learning, & Gender	A+ Alliance for Inclusive Algorithms	Multi-stakehold
The Buenos Aires Declaration	International Telecommunication Union	IO
Artificial intelligence technologies and telecommunications/ information and communication technologies		
THE CORE PRINCIPLES OF THE AI FOR SAFER CHILDREN INITIATIVE	UN Inter-regional Crime and Justice Research Institute	IO
Principles for Responsible AI Innovation		
The Presidio Recommendations on Responsible Generative AI	World Economic Forum	Multi-stakehold
Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy	US Government	Gov't-led
Working Group recommendations for a less fragmented AI policy landscape, incorporating Global South perspectives	Paris Peace Forum	Multi-stakehold
WORKING GROUP ON THE SUSTAINABILITY OF JOURNALISM-KEY RECOMMENDATIONS		
ETHICS GUIDELINES FOR TRUSTWORTHY AI	European Commission	Multilateral
POLICY AND INVESTMENT RECOMMENDATIONS FOR TRUSTWORTHY ARTIFICIAL INTELLIGENCE		
Assessment List for Trustworthy AI (ALTAI)		
NATO AI strategy	North Atlantic Treaty Organization	Multilateral
NATO's Autonomy Implementation Plan		
Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development	United Nations General Assembly	IO
ASEAN Guide on AI Governance and Ethics	ASEAN	IO
Draft Continental Artificial Intelligence Strategy	African Union	IO
AI Safety Governance Framework	China	Gov't-led

Cite this article: Qiu J and Cheng L (2026). Consensus and legitimation in global AI regulations: a sociosemiotic perspective. *International Journal of Law in Context* **22**, 340–361. https://doi.org/10.1017/S1744552326100469

https://doi.org/10.1017/S1744552326100469 Published online by Cambridge University Press