Artificial Intelligence Review (2025) 58:265
https://doi.org/10.1007/s10462-025-11203-z

Web Intelligence (WI) 3.0: in search of a better-connected world to create a future intelligent society

Hongzhi Kuai[1,2,8] · Jimmy X. Huang[3] · Xiaohui Tao[4] · Gabriella Pasi[5] · Yiyu Yao[6] · Jiming Liu[7] · Ning Zhong[8]

Accepted: 23 March 2025 / Published online: 4 June 2025

Abstract

Over the past two decades, Web Intelligence (WI) has emerged as a key field driving the evolution of AI in the connected world, addressing the demands of a future intelligent society. This paper provides a comprehensive review of WI's contributions since its inception in 2000, spanning three distinct phases: Wisdom World Wide Web (WI 1.0, 2000–2009), Wisdom Web of Things (WI 2.0, 2010–2017), and Wisdom Web of Everything (WI 3.0, since 2018). For each phase, we examine key advancements, challenges, and future directions from the perspectives of both intelligent machines and human experts, highlighting significant societal impacts. To advance WI research, we propose a large language model-based learning framework for topic analysis and trend prediction. Moving beyond single-perspective approaches, we emphasize the Connected Intelligence Ecosystem defined by the *HIGH5* scheme comprising one goal, two twins, three fundamentals, four functions, and five services that are realized through WI 3.0. This vision serves as a bridge from localized models to a global reference framework for addressing sustainability challenges in future societies. To illustrate the real-world implications of WI 3.0, we present case studies focusing on brain-inspired research, particularly in the intersection of brain intelligence, brain health, and brainternet-fostering interdisciplinary collaboration across diverse research communities.

Keywords Web Intelligence (WI) · Artificial Intelligence (AI) · Brain Intelligence · Connectedness · Wisdom Web of Everything (W2E) · Future Intelligent Society

1 Introduction

Shaping a more intelligent society remains a shared aspiration of all human beings, with its definition evolving alongside advancements in science and technology. One perspective aligns with the concept of Smart City, where intelligence is embedded into urban spaces to enhance human cognitive capacities, optimize resource utilization, and foster innovation through the integration of information and communication technologies. Another perspective corresponds to Industry 4.0, that is the fourth industrial revolution, characterized by

Extended author information available on the last page of the article cyber-physical systems (CPS) and production processes driven by heterogeneous data and knowledge integration (Lu 2017). From a human-centric standpoint, an intelligent society is envisioned as one where the deep fusion of social, cyber and physical spaces that enables a balance between economic progress and social well-being. This vision emphasizes the delivery of highly personalized goods and services that address diverse potential needs, regardless of locale, age, gender, or language (Deguchi 2020; Jeste 2020).

Despite these varied definitions, interaction and communication still play fundamental roles, requiring advanced connectivity techniques and methodologies. Over the past two decades, Web Intelligence (WI) has pioneered a novel paradigm from the perspective of "AI in the Connected World", laying the groundwork for the emergence of a future intelligent society while shaping its evolving themes and definitions:

WI 1.0: 2000⤳2009★Wisdom World Wide Web

"Web Intelligence 1.0 laid out the blueprint for scientific research and development that explores the fundamental roles as well as practical impacts of Artificial Intelligence (AI) and advanced Information Technology (IT) on the next generation of Internet and World Wide Web (WWW)-empowered products, systems, services, and activities." (Zhong et al. 2000; Yao et al. 2001; Liu et al. 2003; Zhong et al. 2003; Zhong et al. 2007). The main research tracks include Intelligent Web-Based Business; Knowledge Networks and Management; Ubiquitous Computing and Social Intelligence; Intelligent Human-Web Interaction; Web Information Management; Web Information Retrieval; Web Agents; Web Mining and Farming; and Emerging Web Technology and Infrastructure.

WI 2.0: 2010⤳2017★Wisdom Web of Things

"Web Intelligence 2.0 put forward a holistic intelligence methodology and a practical technological way to realize the harmonious symbiosis of humans, computers and things in the hyper social-cyber-physical spaces, through embracing WWW, Web/Internet of Things (W/IoT), as well Information and Communications Technology (ICT)" (Zhong 2013, 2015; Zhong et al. 2016). The main research tracks include Web of People; Web of Trust; Web of Things; Web of Data; Web of Agents; and Emerging Web in Health and Smart Living.

WI 3.0: 2018⤳Onwards★Wisdom Web of Everything

"Web Intelligence 3.0 creates a new vision to build upon thinking-supported techniques and methods used above the social-cyber-physical space, featured by integrated investigations of AI and brain intelligence with big data, a new way from the systematic brain-machine intelligence research to a new wisdom service chain for connected everything" (Kuai et al. 2020; Kuai and Zhong 2020; Kuai et al. 2022). The main research tracks are enriched and extended as Web of People; Web of Trust; Web of Things; Web of Data; Web of Agents; Web in Industry, Society, Health and Smart Living; and Fairness, Accountability, Transparency, and AI Generated Content (AIGC).

It can be seen that the evolution of WI is synchronous with the most advanced achievement in these changing times, under a number of thematic and cross-cutting areas, such as AI, data science, computing abilities, web popular degrees, information and communications technology. To gain a better understanding of the WI field, Fig. 1 provides a com-

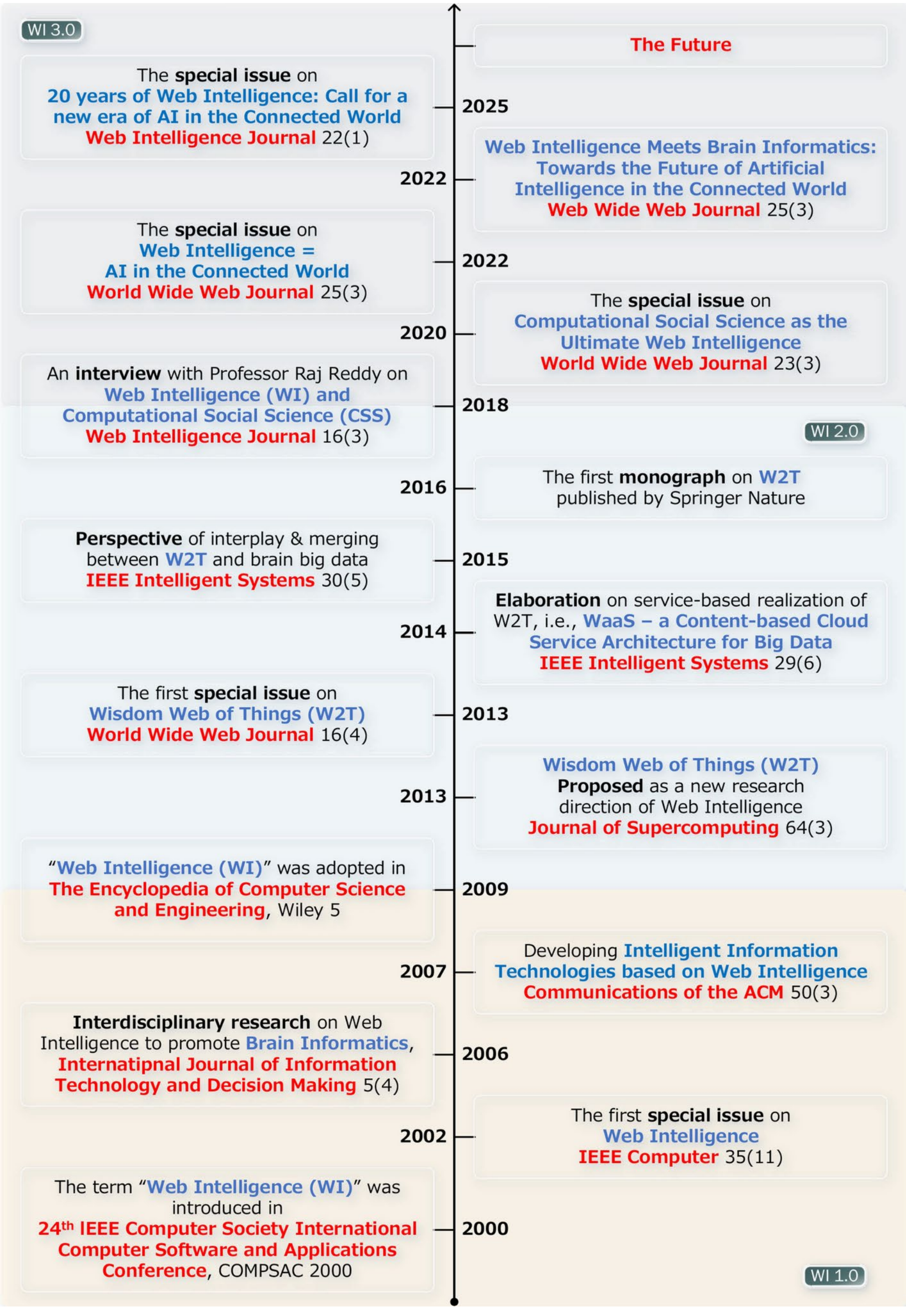

Fig. 1 A 25-year journey of Web Intelligence through the periods of WI 1.0 (Wisdom World Wide Web), WI 2.0 (Wisdom Web of Things), and WI 3.0 (Wisdom Web of Everything)

prehensive overview of representative events of WI from 2000 to future, throughout three stages of WI 1.0, WI 2.0, and WI 3.0.

As any occurrence of new technology and methodology depends on the impact of past era, along with the eager requirements for current and future moment, this paper conducts a comprehensive review of WI, presenting an overview of the content, scopes, and findings on WI by examining existing studies in all databases within the Google Scholar database. Altogether, over 10,000 snippets of papers related to WI are grouped into various research tracks and are analyzed by different pre-trained large language model techniques. Looking ahead, this paper presents the paradigm of WI 3.0 to help build a future intelligent society, revealing the high degree of integration of cyber and physical spaces, the new technology of wisdom services in the connected world promoted by the interconnected symbiosis of "human–machine–thing", new consumption and production patterns, as well as new industries and formats of innovative service-oriented society. Specifically, five research questions (RQs) will be addressed in this study.

RQ1 What are studies and practices related to the scopes of WI 1.0, WI 2.0, and WI 3.0?

RQ2 How can natural language processing and word embedding techniques be integrated to determine the time scale of research trend shifts?

RQ3 How can generative techniques, such as large language models, be applied to explore potential research pathways across different domains and scopes?

RQ4 What are the core schemes of future intelligent society from the WI perspective?

RQ5 How can WI research be integrated into the development of a future intelligent society?

The key contributions of this study are summarized as follows: (1) introduction to WI research, a rapidly evolving field to AI in the connected world, along with a comprehensive review of representative works across WI 1.0, WI 2.0, and WI 3.0; (2) a general framework for analyzing the temporal evolution of research trends, leveraging natural language processing and word embedding techniques combined with similarity measures across multiple time scales; (3) a predictive framework for identifying future research directions by extending past and current topics from both temporal and domain perspectives, integrating generative models and prompt-based analysis; (4) a roadmap outlining prevalent research topics and potential future pathways within the WI ecosystem; (5) an exploration of critical challenges and interoperability issues in WI 3.0, developed in collaboration with researchers from diverse backgrounds; and (6) a series of emerging paradigms within a conceptual framework regarding the interconnected ecosystem towards future intelligent society. Figure 2 illustrates the organization of this paper.

In the rest of the paper, we firstly explore the WI-related studies during different periods, starting from the background to key topics and their relevant works at the end. The time spanned over twenty years, from 2000 to 2025, during which the three periods are reviewed in Sects. 2 and 3, respectively. Challenges and directions for future research are predicted and introduced by combining foundation models in Sect. 4. An ecosystem of interoperability for future intelligent society is proposed as well in Sect. 5. Section 6 summarizes and concludes this paper.

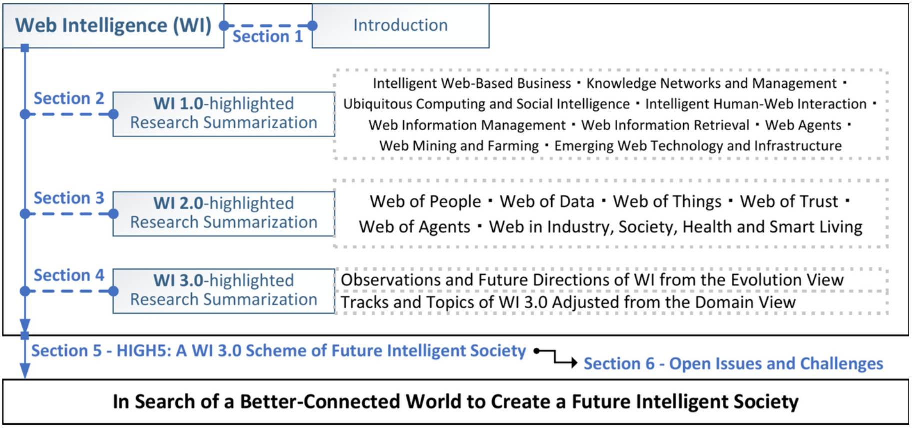

Fig. 2 The organization of this paper

2 WI 1.0: Wisdom World Wide Web-W4

The field of Web Intelligence in the first decade of the 21st century boldly ventured along several major methodological dimensions, as we can readily note in our retrospective research snapshots.

- *Intelligent Web-Based Business*: Levin and Zahavi discussed the effects of data mining on targeting applications from three types of targeting models, including continuous-choice, discrete-choice and in-market timing models (Zhong et al. 2004). Niu et al. (2007) developed a Web-based cognitive business intelligence system through integrating the executive's situation awareness, mental models, and other past experience. Baumgartner et al. (2005) described a solution to retrieve and normalize data from public information sources to structure formats automatically, and further supporting WI applications. The authors developed a WebCompare system to mine the unexpected information from company Web sites (Liu et al. 2003).
- *Knowledge Networks and Management*: The semantic similarity measures were studied in Liu et al. (2007), focusing on named entities in information retrieval and knowledge management. Ontologies, as an important knowledge management tool, have been widely integrated into different web applications, such as ontological user profiles representation (Sieg et al. 2007), product data semantics interoperability (Abdul-Ghafour et al. 2007), text summarization (Hennig et al. 2008), personalized web information gathering (Tao et al. 2010), on-line news recommendation (Cantador et al. 2008), semi-automatic OWL ontology refinement (Völker et al. 2008), taxonomic relation mining from the Web (Maedche et al. 2002), and ontological representation and construction (Zhong 2002). Martin systematically summarized various projects concerning knowledge representation, sharing, and retrieval on the Web, as well as their details requirements (Martin et al. 2003).
- *Ubiquitous Computing and Social Intelligence*: In community detection and its extended scenarios, some solutions and viewpoints were proposed, such as the density-based graph clustering algorithm (Falkowski et al. 2007), mutual awareness expansion based

thematic communities analysis (Lin et al. 2007), and joint learning of both positive within-group relations and negative between-group relations in signed social network (Yang et al. 2007). Nishida (2007) introduced the concept of social intelligence design and its technical aspects to support from different groups and teams to the entire society, as well as their communications. A review was given by Borgatti et al. (2009), focusing on network analysis in the social sciences. Thelwall (2005) introduced scientific Web to realize academic hyperlink analysis from university's Web sites to reveal interconnections between research fields. The similar viewpoint was observed by Fu et al. (2007), which proposed an algorithm to discover experts using social network analysis.

- *Intelligent Human-Web Interaction*: Liu et al. (2003) introduced an adaptive user interface to improve user experience through personalized learning. Xu et al. (2008) proposed a framework from analyzing sequenced web pages to web recommendation using Latent Dirichlet Allocation. Emotion and opinion play a very important role in human–machine interaction processes, and further derive a series of studies, such as identification of opinion leaders (Zhai et al. 2008), quotations analysis (Balahur et al. 2009), emotion classification (Lin et al. 2008). Takama introduced the concept of humatronics with the purposes of establishing the symmetric interaction between human and electronic systems within advanced computer networks (Takama and Saeed 2007). In addition, the satisfied and accepted degrees have impact on the human-Web interactions, which are studied in Liu et al. (2002) and Bradshaw et al. (2004).
- *Web Information Management*: Yang et al. (2003) proposed a Web information-extraction system, namely XTROS, combining the effectiveness of domain knowledge with wrapper induction. Rigutini et al. (2005) proposed an Expectation-Maximization based learning algorithm to organize and manage documents in a multilingual environment. Domingues et al. (2007) proposed a data warehouse for the purposes of managing the services and content available, and making the site dynamically adequate to user's needs. Zorrilla et al. (2005) designed a data e-learning Web-house to support e-learning personalized systems. Tan (2002) described a system, namely Flexible Organizer for Competitive Intelligence, transforming raw URLs returned by internet search engines into personalized information portfolios. Furthermore, fuzzy logic and hybrid approaches were discussed to Web-based intelligence gathering and information management (Berkan and Trubatch 2002). To manage complex information, the details of graph database models were discussed in Angles and Gutierrez (2008), including data structures, query languages, and integrity constraints.
- *Web Information Retrieval*: Speretta et al. (2005) studied the personalized search methods through analyzing individual search activities. Massa et al. (2005) proposed a PagereRank method to provide more trust and precise Web. The topics of Web image retrieval were focused in Lin et al. (2003) using a relevance model. Machine learning approaches were used to rank Web pages, such as graph neural network (Yong et al. 2008). Nie et al. proposed their concerns about cross-language information retrieval (Nie et al. 2002). Stronge et al. explored the effects of strategy use and age on search success (Stronge et al. 2006). Lu et al. (2008) proposed a deep Web crawling method to optimize retrieval satisfies at a low cost. Hoeber gave a systematic summary on the key issues, challenges, and opportunities in Web information retrieval support systems research (Hoeber et al. 2008). The rough set techniques were used to Web search result clustering (Ngo et al. 2005). To an information retrieval issue, some researchers not only

focused on the algorithms themselves, but also the performance evaluation methods (Zhou and Yao 2010).

- *Web Agents*: Liu et al. (2003) characterized user behavior from recorded Web log data through an information foraging agent-based approach based on Web regularities. Some authors studied the different characteristics of autonomous agents and multi-agent systems, such as learning, self-organization, collaboration, and adaptive computation (Liu 2001; Liu et al. 2004). Bryson et al. (2003) introduced agent-based composite services within a Web service ontology, supporting us to design a behavior-oriented intelligent semantic Web. Johnson described a way to improve the quality of Web-based education using agent technology (Johnson et al. 2003). Tan et al. developed methods to detect camouflaging and previously unknown Web robots, not the creation of a robot itself (Zhong et al. 2004).
- *Web Mining and Farming*: Han et al. considered how to bring intelligence into Web interactions through uncovering and cataloging the authoritative links, traversal patterns, semantic structures, the semistructured HTML, XML, and database-service-engine information stored on the Web (Han and Chang 2002). Li et al. (2006) introduced implementations and examples of network thinking and network intelligence from different viewpoints of particular nodes, topological measures, structures, static and dynamic properties. Chikhi et al. (2007) comprehensively compared different dimensionality reduction techniques for web structure mining. Alam et al. (2008) described a particle swarm optimization algorithm for the clustering of Web user sessions. Inui et al. (2008) developed an application and database to mine instances of personal experiences and opinions from user-generated contents, such as Weblog and forum posts. Zhong et al. (1999) presented a theoretical framework for peculiarity-oriented mining to mine interesting rules and patterns from multiple databases. Granular and grid computing strategies were widely used in the processes of data mining and knowledge discovery (Yao 2006).
- *Emerging Web Technology and Infrastructure*: Zhong et al. (2007) highlighted the potential of WI-centered intelligent information technologies on living, working and playing, opening a road towards the next generation information technology society. Bodorik et al. (2008) designed a Web services platform with mechanisms to support privacy policies in the agent-based enterprise information technology infrastructure. Kalles et al. (2003) systematically discussed the algorithms of a Web-based intelligent system spanning from site design to user walk-through and extended user-input analysis. Cercone et al. discussed the effects of computational intelligence on Web's intelligence in some topics of machine translation, machine learning and user interface design (Cercone 2002). Personalization was considered at the levels of systems (Castellano et al. 2008) and content (Tan et al. 2004).

In addition, the important aspects of WI were also discussed in education research, such as the combinations of ontologies, adaptivity, personalization and agents (Devedžić 2004; Butz et al. 2006). The WI and intelligent agent technology were also used in complex urban environments, restaurants, and immunology. In the later stage of WI 1.0, the interplay of ideas between WI and other fields became more frequent, for example bioinformatics and brain informatics, which added a new dimension for the next era of WI.

3 WI 2.0: Wisdom Web of Things-W2T

From 2010 to 2017, more and more results have been presented that enrich the theme of WI in its 2.0 era. Here, we review various endeavors that have studied Web Intelligence and potential applications from these five topics of the Web of People, the Web of Data, the Web of Things, the Web of Trust and the Web of Agents, each of which lists several subtopics.

- *Web of People*: Faraj et al. (2011) proposed the concept of the fluctuations in tensions of a dynamic flow of resources in identities, which supports knowledge collaboration in online communities. Guzdial (2013) introduced a viewpoint of human-centered computing in education by combining computing and psychology. Yuan et al. (2013) developed an online search program for homebuyers-oriented recommendations, improving real estate searches using search behaviors, case-based reasoning and ontological techniques. Different techniques and methods were discussed to sentiment analysis and opinion mining, such as statistics and information fusion, from different levels of document, sentence, aspect, comparative analysis and lexicon acquisition (Feldman 2013; Balazs and Velásquez 2016). Becker et al. (2017) presented theoretical predictions and experimental findings on social influence from social network dynamics. Kassak et al. (2018) gave a survey related to acquisition and modelling of user behavior on the Web from the long- and short-term perspectives.
- *Web of Data*: Data provenance and semantic techniques played an important role on big data integration processes (Tomazela et al. 2013; Knoblock and Szekely 2015; Herschel et al. 2017). For big data analytics, Kambatla et al. (2014) and Tsai et al. (2015) provided an overview of the hardware, software, and application landscape, as well as the future directions. Dhar highlighted the view of predictive modeling in data science to cope with increasingly heterogeneous and unstructured contents generated by humans and computers (Dhar 2013). Jordan and Mitchell (2015) gave future trends, perspectives and prospects of machine learning on algorithms and theory at the core of AI and data science. The data-driven services and applications were studied to assist situational development, such as service computing and the description of service data correlations (Gu et al. 2010; Liu 2014).
- *Web of Things*: Atzori et al. (2010) discussed different paradigms of Internet of Things from "Things-"oriented versions, "Internet-"oriented versions, and "Semantic-"oriented versions, as well as their combination. Yan et al. (2012) gave a survey on smart grad communication infrastructures to serve as more and more entities, with the goal of addressing emerging distributed challenges. Steen et al. comprehensively introduced distributed systems from decentralized computers, machines and mobile devices, with two characteristic features from collection of autonomous computing elements to single coherent system (van Steen and Tanenbaum 2016). In this stage, different types of computing modes were further utilized and refined to cope with online data and physical objects, such as cloud computing (Armbrust 2010), ubiquitous computing (Friedewald and Raabe 2011), mobile edge computing (Mao et al. 2017), and transparent computing (Ren et al. 2017). Furthermore, on the basis of W4, the concept of "Things" was highlighted to extend the boundary of "network" wisdom, opening a new era on W2T (Zhong 2013; Chen 2014; Zhong 2015; Zhong et al. 2016), gaining attention to develop techniques, methods and applications around a cycle of "from things to data, informa-

tion, knowledge, wisdom, services, humans, and then back to things".

- *Web of Trust*: Bannister et al. discussed the benefits of information and communications technology to enhance transparency, with the emergence of e-government (Bannister and Connolly 2011). Meanwhile, many studies had focused on trust management, computational trust and reputation models for open multi-agent systems (Pinyol and Sabater-Mir 2013; Yu et al. 2013). Faroukhi et al. (2016) provided a comprehensive review of value creation, data value, and big data value chains to support data monetization services and applications with potential requirements of data security. Yang et al. (2016) proposed a social collaborative filtering algorithm based on trust for recommendations, specifically for sparse data and cold-start users. Ding et al. gave a review on multivariate public-key cryptosystems and multivariate schemes in digital signatures and public-key encryption as well as their security (Ding and Petzoldt 2017).
- *Web of Agents*: Liu et al. (2011) proposed a decentralized, self-organized algorithm, simulated by autonomous multi-entity systems, for solving combinatorial optimization problems. McRorie et al. (2012) discussed the development and evaluation of virtual agents based on sound psychological principles, and the effects of personality on agents' behaviors and productions. Local and global interactions in agent networks were studied to sequential state estimation and small-world social network generation with dynamics (Hlinka et al. 2013). For agent-based modelling and simulation, Macal (2016) gave a systematic review.

In this stage, a track of "Emerging Web in Health and Smart Living" starts paying more attention for scientists from different backgrounds, owing to closer integration of life science and AI. A series of works were proposed to improve the effects of Web and Internet in the scenarios of health and smart living. For example, Yang et al. (2016) proposed an epidemic-model-based tensor deconvolution framework to analyze the spatiotemporal patterns of social contact for understanding and predicting the prevalence of infectious diseases. Vlacheas et al. (2013) proposed a cognitive management framework for IoT to substantially support sustainable development of future smart cities. For the elderly and disabled people, Hussain and Rafferty et al. developed health and emergency-care platform, as well as the intention recognition approaches (Hussain et al. 2015; Rafferty et al. 2017). Facing the increasing big data resource, Pramanik et al. (2017) proposed a conceptual model on big data enabled smart healthcare system framework.

4 WI 3.0: Wisdom Web of Everything-W2E

4.1 Investigative strategies and views

As the continuation of WI 1.0 and WI 2.0, WI 3.0 extends and enriches their topics, fostering new meanings. To systematically reveal the WI's journal during the past twenty years and give its future trending prediction, a corpus was constructed firstly. It includes the focused fields-centric text selected from different sources, where the searched phrase is "Web Intelligence" that is observed by the Google Scholar engine, between 2000 and 2023. Furthermore, we screened the satisfactory text data from the corpus, followed by the PRISMA guidelines (Page 2021). A detailed text selection pipeline is given in Fig. 3, containing the WI-related text data that are used in this paper.

Coupled with these selected texts, a foundation models-fused learning framework (Fu3ML) is proposed to achieve the functions of automatic domain review and quantitative trending predictions in Fig. 4, including the following three components:

- the ***entity recognition component (EnRC)*** is constructed to extract keywords from unstructured text, which can be perceived as evidence to help us clarify a field's development;
- the ***content generating component (CoGC)*** is constructed to create new content based on prompts, which can be identified as inspirations and cues for the future works;
- the ***embedding learning component (EmLC)*** is constructed to learn text representations, which can be used to execute text similarity evaluation and topic cluster learning

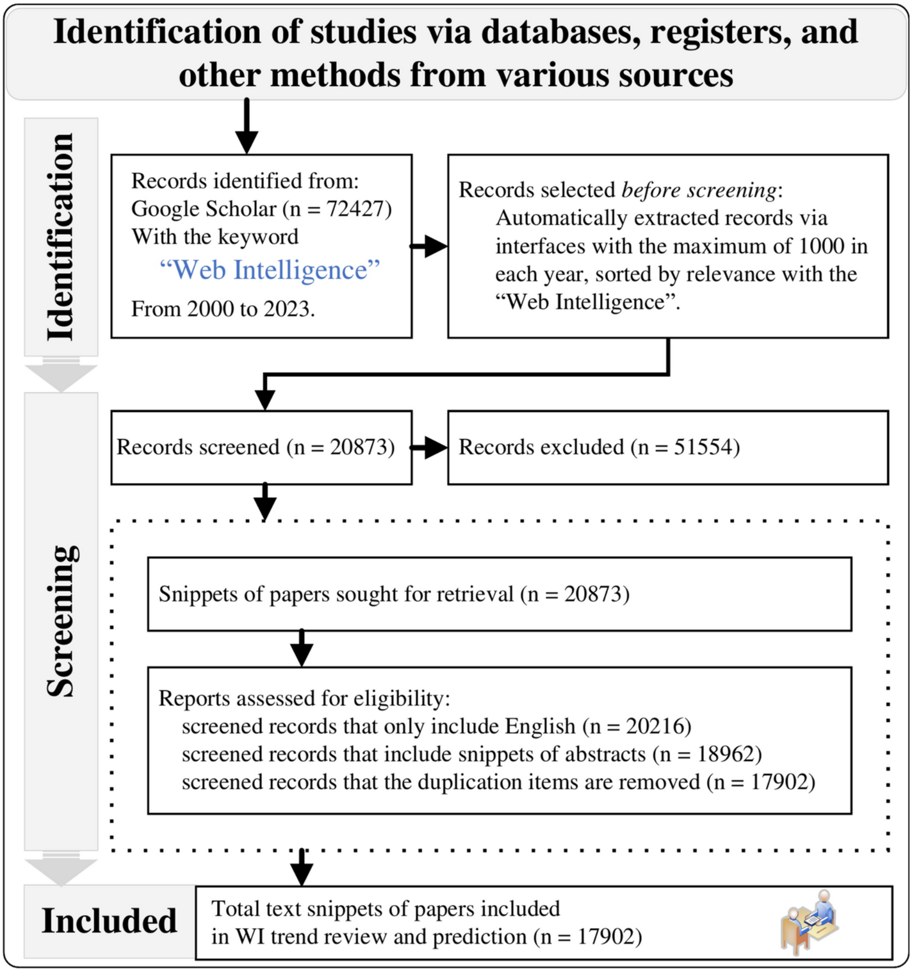

Fig. 3 Flow diagram obtained from the PRISMA guidelines for text selection and corpus construction (Page 2021)

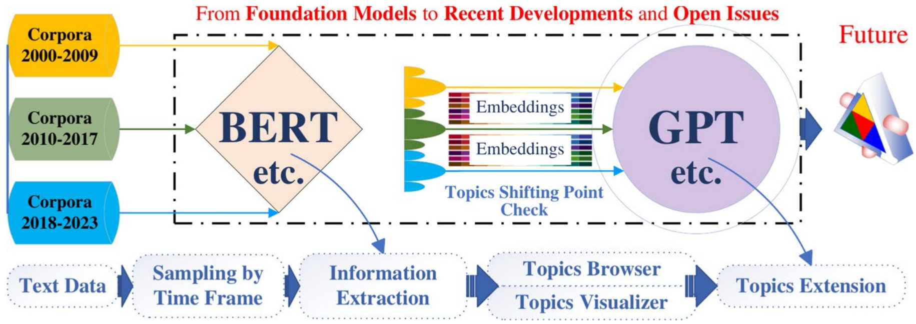

Fig. 4 The topic learning framework based on pre-trained language models and a topic switching point-oriented detection and prediction method

for obtaining a more representative trend;

where such foundation models with different characteristics can be systematically integrated in an unified learning framework, namely Fu3ML, such as a BERT model (Mohammed and Ali 2021) used in the ***entity recognition component***, a GPT model (Liu et al. 2023) used in the ***content generating component***, and a ConSERT framework (Yan et al. 2021) used in the ***embedding learning component***. Through making full use of their respective advantages, Fu3ML can flexibly exchange and integrate different pre-trained foundation models, and further supports downstream developments and complex open issues.

As a typical downstream application, the *EnRC* in Fu3ML is used to recognize the entities from the selected WI-related text, the results of which are counted to analyze research activities during a given time period and the future trends. In particular, we investigate the WI-related field's history and future from the time evolution and domain views. On the one hand, the automatically recognized entities are separated to different groups on the basis of different time periods. For example, the word clouds are created corresponding to the three stages of WI 1.0 (2000–2009), WI 2.0 (2010–2017) and WI 3.0 (since 2018) in Fig. 5a, respectively. It can be found that the degree of interest changes throughout the whole development stage of WI research, such as the system related directions with a sustained focus while the data and user related directions with an increased focus. On the other hand, the hand-crafted entities are created with different tracks and topics as shown in the "Evolution of WI Tracks and Topics" sections of Appendix 1. Therefore, both of them can be compared and verified within and between solutions. Facing an open issue of predicting the future directions and trends in WI, the *CoGC* in Fu3ML helps us generate new contents with respective to different time steps and tracks through prompts-based technologies. Herein, the prompts are constructed with time and scope characteristics from the WI-related entities created by both manual and automated configuration modes. Furthermore, the *EmLC* in Fu3ML is used to encode collected topics as vectors, and then helps us identify the representative topics throughout the whole above-mentioned operations, with detailed experiment settings in Appendix 3. Consequently, the results of trend summarization are obtained as shown in Figs. 8, 9, 10, 11, 12, 13, 14 of Appendix 4; the results of the trend shift degrees are obtained as shown in Fig. 15 of Appendix 4; and the results of trend prediction are obtained as shown in Figs. 16 and 17 of Appendix 4. In the next subsections, Fu3ML will be used to explore WI 3.0 from temporal and domain views.

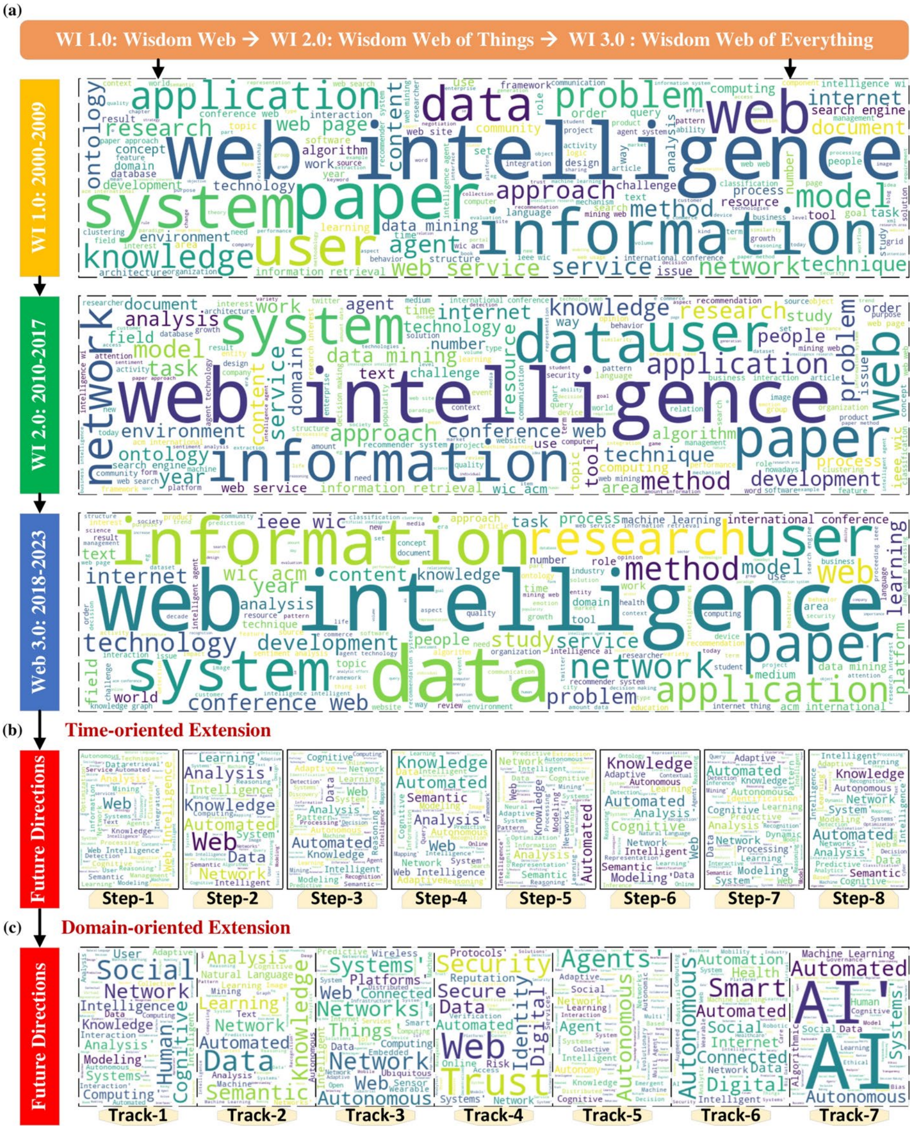

Fig. 5 Illustration of topics distributions from historical and perspective views. **a** The research topics are reviewed at different stages of WI 1.0 (2000–2009), WI 2.0 (2010–2017), and WI 3.0 (since 2018); **b** the future directions are predicted along with the time dimension from Step-1 to Step-8; **c** the future directions are predicted along with the domain dimension from Track-1 to Track-7

4.2 Exploring the WI 3.0 from the temporal view

Similar as three waves of AI evolution with significantly different research trends, some changes also can be observed during progress of WI. One of the most interesting viewpoints is how to get the switching points of trend changes via automatic text mining methods. To be honest, it is difficult to give clear boundaries to obviously split different phases, because the development of technology and methodology is smooth with the continuous and overlapping characteristics. It means that the task of capturing the finer and temporal research topic changes is not easy, but we believe it is worth to pay more attention and make additional efforts. For this purpose, we propose a topic switching point-oriented detection and prediction method to identify and trace the topic development and changes in Appendix 2, with a short name of TSPDP.

Accordingly, the topic trend prediction results are given from the incremental view with the time window of four years in Fig. 5b, where the time window is calculated in Algorithm 1 of Appendix 2 and is identified in Fig. 15 of Appendix 4. It can be found that the size of topic diversity indicates the periodic characteristic, through observing the intensity of keywords in word clouds. In particular, the topic diversities at step–1, –3, –5 and –7 are bigger than those of diversities at step–2, –4, –6 and –8, where the former implies more research directions than the latter. It is consistent with the law of scientific development, growing in fluctuations. Furthermore, the Web-related fields remain play a major role in the early stage, while the human intelligence and cognition-related subjects are observed with increasing weightings in the near future, through observing the detailed topic distributions in Fig. 16 of Appendix 4. It responds to the future trends toward embodied AI. On the basis of the current results, it suggests us to put more efforts on the AI fields such as adaptive agents, autonomous systems, semantic informatics, coupled with the brain intelligence fields on cognition, thinking, reasoning, learning, decision, emotion, and so forth.

4.3 Exploring the WI 3.0 from the domain view

In this part, the research trends of WI are predicted from different domain views. Different the trend prediction from the temporal view, depending on the multiple iteration learning and automatic collected topics, multiple research trends are extended once at different smaller domains in this section, where the prompts include the hand-crafted topics that are created by experts. Therefore, the breadth of WI is explored from the seven domains with respect to the following seven tracks.

Track-1: Web of People;
Track-2: Web of Data;
Track-3: Web of Things;
Track-4: Web of Trust;
Track-5: Web of Agents;
Track-6: Web in Industry, Society, Health and Smart Living; and
Track-7: Fairness, Accountability, Transparency, and AI Generated Content.

In details, the *CoGC* in Fu3ML is used to perform the task via Algorithm 2 of Appendix 2, extending new contents based on the given keywords (see the "Tracks and Topics of WI 3.0" section of Appendix 1).

As a result, the core extended topic trends are summarized corresponding to the given seven tracks as follows, which are also visualized by word clouds in Fig. 5c. It can be found that the diversities of extended topics in different tracks changes, depending on the size of prompts. Meanwhile, these new generated topics demonstrate relatively clear boundaries. Furthermore, the detailed statistics of topic distributions are given in Fig. 17 of Appendix 4: for Track-1, these topics are highlighted such as collective intelligence, human computing, human behavioral analysis, user interfaces, and cognitive social research; for Track-2, these topics are highlighted such as data mining, knowledge representation, information extraction, natural language processing and generation, and big data studies; for Track-3, these topics are highlighted such as interconnected network, Web, sensor and things, coupled with cloud, edge and quantum computing; for Track-4, these topics are highlighted such as trust, security, verification, reputation, identity, risk, and privacy; for Track-5, these topics are highlighted such as autonomous agents, intelligent systems, symbolic agents, cognitive agents, reinforcement learning agents, decision support systems, and evolutionary agents; for Track-6, these topics are highlighted such as automation in industry, intelligent logistics, data-driven smart cities and utilities, virtualized society, wearable and digital health; for Track-7, these topics are highlighted such as the AI research on autonomy, interoperability, accountability, traceability, explainability, ethics, transparency, bias, governance, and risk awareness and regulation.

As demonstrated above, several open research issues have been identified to put forward for the further evolution of WI 3.0, AI in the connected world. We believe that these findings will stimulate future research, bridging the cross-cutting areas of AI, brain intelligence and connectedness, by raising awareness of each other and pushing forward new topics to create in future research endeavors.

4.4 Embracing different domains with WI 3.0

Facing the development of future WI 3.0, we first collect and summary the research works distributed in different tracks, in collaboration with scientists from different backgrounds such as AI, cognitive science, behavior science, brain informatics, as well internet and communication technology.

- *Web of People*: Gao et al. (2019) discussed the information diffusion of community structure in social networks. Camacho et al. (2020) proposed a set of new metrics for social network analysis from four dimensions, including pattern and knowledge discovery, information fusion and integration, scalability, and visualization. Koukaras et al. (2020) proposed and evaluated a taxonomy of social media types using a hypothesis-based data driven methodology, which could help us capture emerging services in social, entertainment and profiling networks. The authors discussed a diverse range of social groups centered on network structures (Ritchie 2020). In 2017, the Turing Award winner, Prof. Raj Reddy, gave an interview of the effects of Web Intelligence on computational social science (Zhong et al. 2018).
- *Web of Data*: Lnenicka et al. discussed the theoretical background and essential elements in big and open linked data analytics ecosystem (Lnenicka and Komarkova 2019). Heath and Bizer (2022) introduced the concepts and techniques in the field of linked data, evolving the web into a global data space. Ji et al. (2021) gave a survey on

knowledge graphs related to representation, acquisition, and applications. Zhang (2023) and Xia et al. (2024) investigated and discussed the multi-scale data fusion and multi-granularity data computing problems. In the field of information retrieval, Keyvan and Huang (2023) gave a survey of techniques, approaches, and tools to approach ambiguous queries. Huang and his colleagues also focused on individual issues, such as building a topic-graph probabilistic model for web search personalization (Zhao et al. 2022). Some product search models were constructed to assist users in locating desirable items via representation learning (Zou et al. 2023; Bassani and Pasi 2022). The visualization issues based on interactive Web were considered in Sievert (2020) and Postma and Goedhart (2019).

- *Web of Things*: Yue et al. gave a comprehensive survey on the reliability of mobile wireless sensor networks, with the development and popularization of mobile terminal technology (Yue and He 2018). Mahmud et al. (2018) highlighted the effects of fog computing to support the computational demand of real-time latency-sensitive applications of largely geo-distributed IoT devices/sensors. Cao (2022) discussed the conceptual map, research issues, and technical opportunities of decentralized AI and edge intelligence, and their impact on smart blockchain, Web3, metaverse and decentralized science.
- *Web of Trust*: From the technology-oriented trust perspective, Zheng et al. (2018) discussed the challenges and opportunities on the blockchain applications. Haghighi et al. (2020) provided a trust-based public key management solution for linking edge devices in a fog architecture. From the content-oriented trust perspective, Tolosana et al. (2020) reviewed the technology and history of Deepfakes, and its applications on face manipulation and fake detection. Cheng et al. (2019) investigated the social commerce from an integrated view of particularized trust. Kong et al. (2020) examined the effect of social and technical enablers on consumer trust in the context of sharing commerce. To promote responsible trust in AI, Liao et al. (2022) developed a conceptual model called MATCH with a use case and future directions.
- *Web of Agents*: Self-driving vehicles can be regarded as the agents organized into the perception system and the decision-making system in real world, which are explored in Badue (2021). Elshan et al. (2022) conducted a systematic literature review to interpret the combination of empirical knowledge on user interaction with intelligent agents. Xi et al. (2023) discussed the potential of large language model based agents related to brain, perception, and action. In the book, Lugrin et al. (2022) provided a comprehensive overview of the research fields of embodied conversational agents, intelligent virtual agents, and social robotics.

With the increasing of requirements, the tracks of WI are further expended with some new directions of "Web in Industry, Society, Health and Smart Living" and "FaccT and AIGC". For example, under those tracks, Sisinni et al. (2018) gave a systematic review of the challenges, opportunities, and directions on industrial Internet of Things. Qu proposed a blockchained federated learning framework through mixing blockchain technology with cognitive computing, as well integrating the thought pattern of the human mind in industry network (Qu et al. 2020). Tao et al. (2020, 2021) studied health risk prediction through knowledge graph techniques and the radio frequency identification technology with machine learning for remote patient monitoring. Yao introduced a conceptual model and concept learning paradigm to help interpret granular computing in the cognitive era, empowering to push forward the progress of society by constructing a symbols-meaning-value space and applying the principles of three-way decision (Yao 2018, 2022). The authors reviewed the developments and challenges of contact tracing technologies for epidemic prevention and control from both “static and dynamic” and “individual and group” perspectives, especially for data-driven strategies and methods (Chen et al. 2018). The influence of advanced machine learning and deep learning was discussed to provide powerful solutions in various healthcare scenarios (Yu et al. 2021; Chen 2024). The advanced information and communications technology, represented by 5G technology, has begun with a new wave of “pervasive connected world” (Saghezchi et al. 2015). Furthermore, the emerging brain computing techniques and methods accelerate the process of making everything intelligent interconnections and interactions in the connected world, such as brain-oriented analysis, brain-inspired approach, and brain-computer interfaces (Gao et al. 2021; Mehonic and Kenyon 2022; Wang 2024). Meanwhile, the ethical issues have been studied in the development and use of brain-inspired AI (Farisco 2024; Ortega-Bolaños 2024). It is worth noting that the interconnections of those techniques and methods mentioned above can strengthen or substitute human peripheral working capacity, and motivate potential applications in various scenarios of living, learning and clinical towards future intelligence society.

5 *HIGH5*: a WI 3.0 scheme for the future intelligent society

5.1 Emerging paradigms for an intelligent society

Imagine a human-centric society empowered by advanced intelligence, defined as: “An intelligent society is a human-centric, data-driven, and innovative service-oriented society, realized through human–machine co-intelligence, collaborative evolution, and interconnected symbiosis within a connected world.”. It includes various carriers of city, community and people with different intelligent elements under the social-cyber-physical space. From the vision of a future intelligent city, it integrates digital and intelligent twins technology to optimize city services and connect to all societal elements. However, from the vision of an intelligent community, it further strengthens social connection, supporting rich and meaningful cultures as well. Meanwhile, the human beings in the future society will experience the enhanced quality, performance, and interactivity of living services, with reduced costs and resource consumption.

To meet the requirements of creating an intelligent society and make the most of prospective opportunities, we propose a comprehensive scheme, namely *HIGH5* with “one goal, two twins, three fundamentals, four functions and five services” along with the WI 3.0 era. Specifically, one goal is in search of better-connected world to create a future intelligent society. Two twins are digital twin and intelligent twin, fusing brain-machine intelligence via the extensional representation of brain intelligence with big data. Three fundamentals are **A**rtificial Intelligence, **B**rain Intelligence, and **C**onnectedness, namely **ABC**. Four functions in WI 3.0 are recognition, decision-making, discovery, and generation. Five services are provided, such as system as a service, content as a service, wisdom as a service, model as a service, entity as a service in the connected world. Next, more details will be given.

Paradigm one: An intelligent society is aspired that individuals, things and systems are all connected in cyberspace and optimal results obtained by AI exceeding the capabilities of humans are fed back to physical space. One significant feature is a human-centered society that balances economic advancement with the resolution of social problems by a system that highly integrates cyberspace and physical space.

Paradigm two: Two-dimensional twins on digit and intelligence pursue techniques and methods of digital transformation and intelligent decoding and encoding. At the basic level, physical objects in real environments are copied with digital objects in virtual environments. At the higher level, intelligent mechanisms are decoded in physical space, and encoded in cyberspace, achieving the wisdom's transformation. For example, intelligence is transformed from systematic brain-machine intelligence research to new AI industry chain in the connected world.

Paradigm three: Pushing forward three fundamentals of the future intelligent society based on the ABC paradigm, named by three first letters of AI, brain intelligence and connectedness. For example, AI is boomed on generative techniques and methods, machine learning, natural language processing, and so forth. Brain intelligence is processed by Brain Informatics, big brain data science, brain-inspired computing, and so forth. Connectedness focuses on Web of everything, social-cyber-physical spaces, interaction and interface, as well as human–machine symbiosis in a connected world.

Paradigm four: The four major functions are tracking, including recognition, decision-making, discovery, and generation. Through the recognition ability, AI systems can recognize patterns, objects, or data within various forms of input, such as images, text, or speech. The recognition ability is used in image recognition, speech recognition, and conversational scenarios. Through the decision-making ability, AI systems can make informed decisions based on data analysis and predefined rules. In applications like autonomous vehicles and recommendation systems, AI aids in decision-making by processing real-time information. Through the discovery ability, AI is employed to discover hidden insights or trends within vast datasets. Machine learning and data analysis techniques uncover valuable information, aiding in research, business intelligence, and scientific discoveries. Through the generation ability, AI is capable of generating content, such as text, images, music, and videos. Generative AI models like GPT series can create human-like multimedia content and have applications in content creation and creative industries.

Paradigm five: Human-oriented services are provided through executing satisfactory request-response actions interacted with the appropriate objects at a suitable time and context. Those services include, but not limited to, system-as-a-service (such as software, platform, and infrastructure), content-as-a-service (such as data, information, and knowledge), and other viewpoints such as wisdom, model, health, robot, agent, thinking, and entity. Those services are provided in response to the whole lifecycle of human, systems, and things in both virtual and real world.

Coupled with the five paradigms of *HIGH5*, various societal elements can be allocated and integrated to provide personalized services and meet personalized requirements within an intelligent ecosystem, systematically and purposefully.

5.2 Future society empowered by intelligence ecosystem

An intelligent ecosystem was constructed to align with the above paradigms on creating such a future society with advanced intelligence and sustainability in the connected human–machine–things world. As shown in Fig. 6, it mainly includes seven layers: the base layer, the resource layer, the driver layer, the technique layer, the protocol layer, the innovation layer and the service layer. The following attempts to provide a complete solution with respect to different layers.

At the base layer, all objects from physical world are aligned with the theoretical concepts with profile and characteristics specified in society, individually or collectively. Such objects are classified into different working fields, such as people, living, economy, environment, governance, and mobility. Meanwhile, some key factors need to be concerned in response to these fields, such as: the people's factors of social and ethnic plurality, flexibility, creativity, open-mindedness, and participation in individual and public life; the living factors of culture, health conditions, safety, housing quality, education and entertainment; the economic factors of innovation, entrepreneurship, trademarks, productivity, marketing

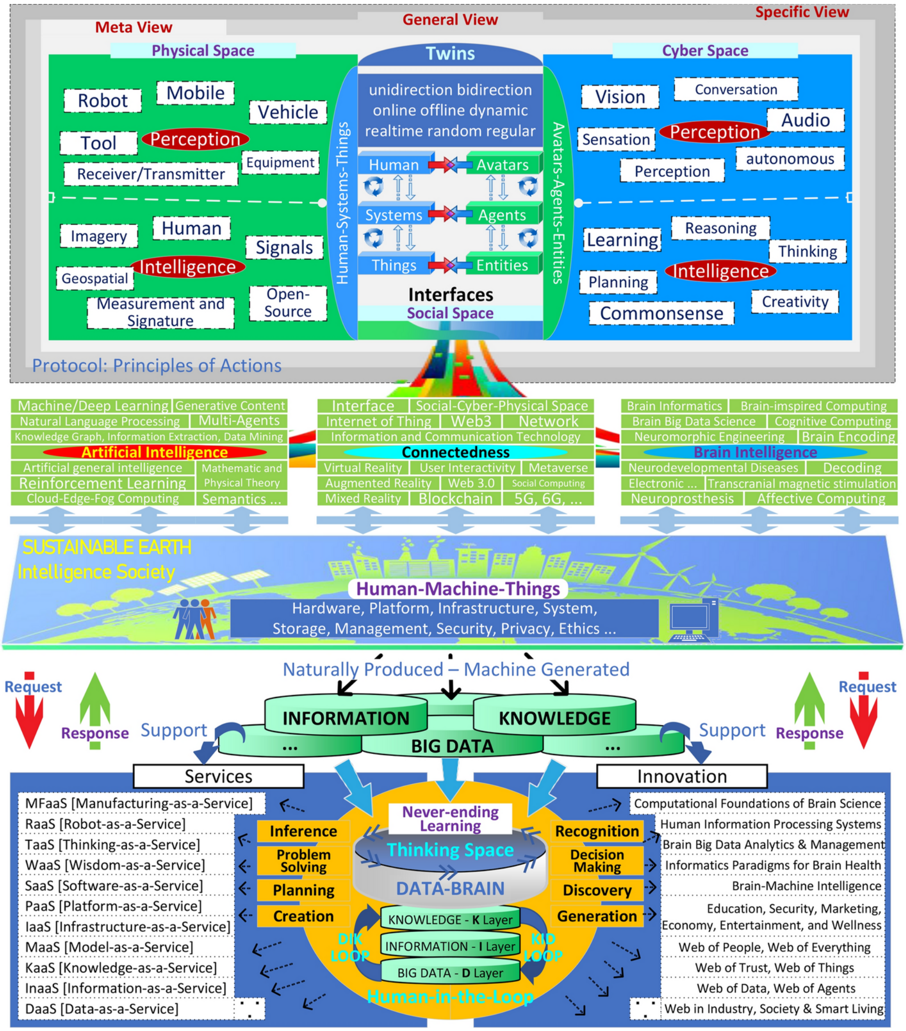

Fig. 6 Envisioning a collaborative intelligence ecosystem framework in search of a better-connected world to create a future intelligent society

and international cooperation; the governance's factors of decision-making, public services, transparent scale, political strategies and perspectives; the environmental factors of natural conditions, pollution reduction, sustainable resource management, and protection; the mobility's factors of local and global accessibility, information and communication infrastructure, and safe transportation.

At the resource layer, the generated rate of resources from human, machine and things in real and virtual world are increasing with accelerating trends at different scales of data, information and knowledge. On the one hand, behavior and activities of objects in the natural world are captured and stored in techniques with stronger digital and perpetual abilities, from which huge number of primary sources are obtained. On the other hand, supported by generative technologies, huge number of machine-generated contents will be posted more and more. Facing these resources, the capabilities of effectively discriminating fake and generated contents should be built firstly. The former focuses on whether one resource can be used or not, while the latter focuses on whether one resource is valuable to be used or not. Furthermore, these resources are systematically organized into a unified one (such as knowledge graph, symbolic representation, and ontology) for ensuring efficiently resource access and computing, where any resource object can know itself and interacts others easily when they are used.

At the driver layer, thinking space is constructed at the top of the resource layer to give wisdom actions, where methods, systems and models are developed via intelligently integrating different resources and techniques. In particular, the following four capacities are developed, including recognition, decision-making, discovery and generation. The thinking is already discussed in an intelligent model, namely Data-Brain, mastering never-ending learning and human-in-the-loop mechanisms that are driven by a loop depended on a three-tier architecture of data, information, and knowledge.

At the technique layer, the means and tools are clustered around three focused topics of AI, brain intelligence and connectedness. For example, the AI part focuses on techniques, such as methods of deep, machine, generative and reinforcement learning, data storage and mining, cloud-edge-fog computing, as well as natural language processing and computer vision. The brain intelligence-related techniques include brain-inspired computing and chips, neuromorphic engineering, as well as smart health and living etc. The connectedness part focuses on resource interoperability and sharing, network and mobile communication, distributed/federated learning, user interactivity, extended reality, social cognition, connected vehicles, interactive interfaces, IoT and robotics, and so forth.

At the protocol layer, the principles guide us to form a better social-cyber-physical space from meta, general and specific views, serving the future intelligence society. First, these protocols are formulated at different social-cyber-physical sides, respectively, with respect to perceptual and intelligent requirements. The entities from both physical and cyber sides are connected through the digital and intelligent twins. Both of them share the similar characteristics, such as construction of object's profiles, belonging to the mirror relations. Meanwhile, they also have subtle differences, for example, giving greater consideration to hardware techniques in the physical side and virtual techniques in the cyber side. In the social side, the interactive interfaces of human, systems, and things are constructed with (un/uni/bi/multi)directional, (on/off)line, dynamic, realtime, random, and regular characteristics. In addition, these points should be focused in everywhere, such as security, privacy, trust and accountability.

At the service layer, the requirements of personal services are increasing to balance distribution of resources and improve the satisfaction of all participants for providers and users. For example, scientists require personal data, information and knowledge services in the resource side, as well as personal infrastructure, platform, and software services in the computing side from cloud-based providers; factories require the better services from intelligence manufacturing abilities; and all of them are integrated to provide thinking-supported wisdom services for users.

At the innovation layer, a way is explored to catalyze new products, ideas, and even ways of life from a network of actors, activities, and artifacts, such as individuals, entities, resources, and structures. The innovation is created through different national, regional, and sectoral modes during processes of interaction, cooperation, competition, and selection.

5.3 A roadmap from intelligence research to future intelligence society

Although it is difficult to give a solid definition for the nature of intelligence so far, many attempts have been executed to interpret it from different views of multiscale brain computing, especially for systematic methods like omics. In the few past year, a systems-centric methodology, namely Brain Informatics, is proposed to handle human brain-centric macro-meso-micro big data, individually or integratedly. Brain Informatics treats human brain as an information processing system with big data to perform systematic studies in the full scale data of human brain in vivo with cognitive tasks. The ultimate goal is to seek a complete comprehension of the human consciousness and intelligence.

Inspired by the Brain Informatics methodology, an WI-centric intelligent health technology, namely Data-Brain based general intelligence model, is studied to assist systematic investigation of complex brain science problems (Kuai and Zhong 2020). Data-Brain is constructed in the thinking space to integrate multi-task, multi-modal, multi-scale brain big data as extensional representation of brain intelligence, meeting the requirements of systematic design of cognitive experiments, systematic data management and organization, and systematic data analysis/simulation. The model is a never-ending learning (NEL) agent driven by the data-information-knowledge-wisdom loop with rule-based reasoning and evidence combination with fusion computing (Kuai 2021, 2024). Apart from supporting the study of brain intelligence, such a model has been discussed in the directions of brain health and braininternet from previous studies (Kuai et al. 2022). As a representative case, Fig. 7 gives a specific view of Data-Brain acting on the smart living scenario.

In Fig. 7, a generalized root node of Data-Brain is built in the basic layer, organized by the joint units of data, algorithm and computility. From the data units, the multimodal and multiscale brain big data are systematically collected and managed in a data center, together with multi-granularity information and multidimensional knowledge representation; from the algorithm units, the autonomous and never-ending learning methods are activated to extract enormous information and knowledge from the data continuously; from the computility units, the new computer architecture and systems are developed to meet the personal requirements from the data and algorithms, such as cognitive, molecular, and quantum computers. Facing a specific application requirement on the integration of “health-medical-welfare”, domain knowledge is reorganized from the generalized root node to various personalized service nodes, like the major diseases (such as cancer, heart disease, mental illness, and cerebrovascular disease) in the world. For example, personalized services can

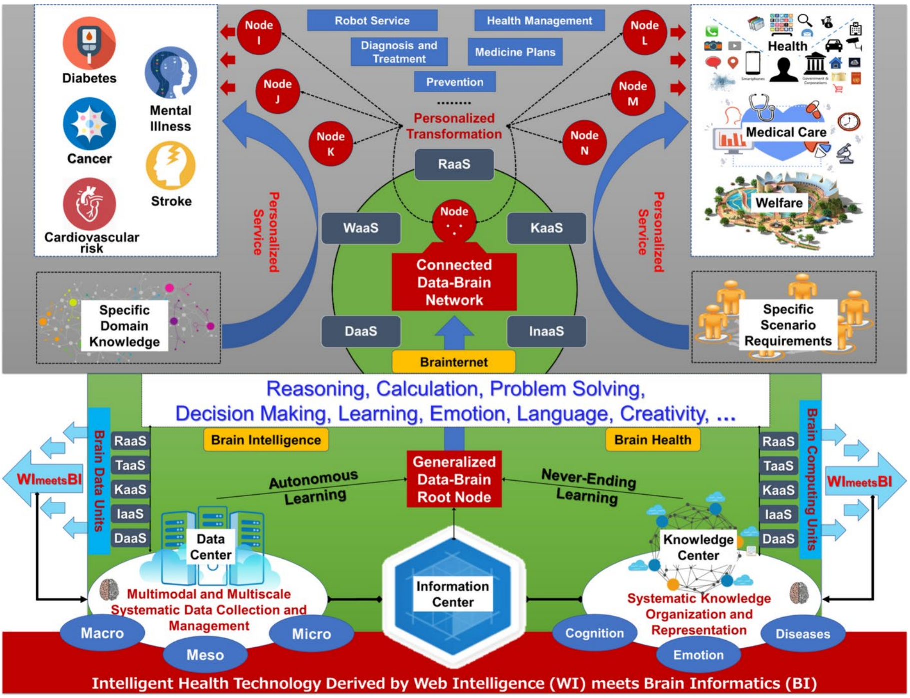

Fig. 7 The smart health practices based on the WI-centric intelligent health technology. *DaaS*, Data as a Service; *InaaS*, Information as a Service; *KaaS*, Knowledge as a Service; *TaaS*, Thinking as a Service; *RaaS*, Robot as a Service

occur in the nodes of health management, medicine plans, prevention, diagnosis, and treatment to realize such a systematic response by enhanced and more effective use of connected Data-Brain network for coordination and cooperation.

During the practical process, the findings from the brain intelligence study can accelerate the progress of intelligent health technology of WI. As a control object, the results of the brain intelligence study can be regards as a reference object to help us decode abnormal mechanisms of brain and mental disorders. Reversely, the abnormal brain studies also can help us resolve the brain information processing mechanisms more in depth, corresponding to different cognitive conditions. Brainternet focuses on the closer combination of brain and web to provide personal services. By meeting Web Intelligence and Brain Informatics (WImeetsBI), the convergence of brain intelligence, brain health and brainternet is one of most important roads towards the future intelligence society. In particular, during the process of creating a future intelligent society, the following open issues and directions have been highlighted, when combining within the ABC paradigm of *HIGH5*.

First of all, bringing together the needs of next-generation AI and society, it will stimulate a constant endeavor to overcome theoretical and practical difficulties as follows:

- What are the key breakthroughs and applications of generative AI?
- What ethical considerations should be taken into account when deploying generative AI systems?

- How can generative AI be harnessed for positive societal impact while mitigating potential risks?
- What role does regulation play in shaping the responsible use of generative AI?
- How can the research community and industry collaborate to ensure the responsible development and deployment of generative AI technologies?
- How can generative AI transform education?
- How can generative AI empower WI to create the future intelligent society?

Furthermore, from the brain intelligence view, the following issues are eager to be resolved, which are considered from other related directions of smart health:

- How can we understand brain intelligence in depth, supported by the investigations of healthy and abnormal brains?
- How can we explore neural mechanisms of cognition, emotion and disease, as well as their biomarkers?
- How can we investigate the disease progression across the whole cycle with respect to the prevention, diagnosis, evaluation, treatment, prognosis and rehabilitation?

In addition, from the connectedness perspective of intelligence and network, the following issues are addressed to promote smart living and brainternet:

- How can we understand brain from neural microcircuits to macroscale intelligence systems, supported by connecting network and brain with big data?
- How can we utilize the power of human brains and man-made networks to create a better-connected world in the 5G era?
- How can we realize human-level collective intelligence as a big data sharing mind on the WI 3.0 by developing brain inspired intelligent technologies to provide wisdom services?

Through the joint methods of qualitative and quantitative evaluation used in this study, WI has demonstrated a path in search of a better-connected world to create a future intelligent society. In particular, the related issues have been discussed during the three significant stages of WI 1.0, WI 2.0, and WI 3.0. Their studies and practices related to different WI scopes have been given corresponding to RQ1, which provide the strategies to face challenges in different fields. A four year-surrounded topic trend switching scale has been identified in the WI field, which meets the rapid development rates on current digital era corresponding to RQ2 and RQ3. It is noteworthy that the algorithm for identifying topic trend switching scale is not limited to the study of WI field. Instead, it can be expanded and used widely to other fields with super, sub, and same levels, such as AI, brain, and computational science. Driven by the practices of WI, we formalize the *HIGH5* scheme, depicting general characteristics of the future intelligent society corresponding to RQ4. In details, we design a connected intelligence ecosystem framework to empower the progress of the intelligent society, bridging the gaps between intelligence research and intelligence practices corresponding to RQ5.

6 Conclusions

Entering the WI 3.0 era, the Web Intelligence field continues to serve as a hub for the cutting-edge research and innovation of AI in the connected world to create a future intelligent society. Especially, the rapid advancements in generative capability and the creativity study in the brain have unlocked a multitude of opportunities, from enhancing creative processes and content generation to addressing complex real-world challenges in industry, finance, culture, education, healthcare, smart city, and more. With these remarkable points, AI Generated Content sheds light on such critical aspects and offer a balanced perspective on Generative AI as well as insights into its capabilities, ethical considerations, and potential societal impacts. In particular, it is necessary to comply with any ethical principle in developing the techniques and methods of WI 3.0, including ethics, moral awareness, consciousness, the problem of opacity, singularity and value alignment, autonomous systems, and other debates.

Generally, the major ethical issues for human societies AI poses are presented in bias, gender, social, racist and sexist, with the goal of regulating an agent's behaviors by developing an ethical dimension. Meanwhile, the moral status of intelligent machines is increasing the discussion of moral and legal rights, containing the sub-discipline technologies of robot and computer ethics. For this, a WI practitioner should align a system with human moral decisions through a theory-driven top-down approach, a bottom-up approach shaped by evolution and learning, and their mixed approaches, inspired by autonomy, indirect duties and social relation. When machines become more and more sophisticated and intelligent, a conscious and living machine is developed by scientists, experiencing pleasure and pain with minds. Along with the intelligence explosion, technological singularity will arrive by creating self-replicating and super-intelligent AI machines to realize the value alignment, and being concerned with humanity as well. To understand the underlying reasons for sorts of decisions, a WI practitioner should follow the principles of explainability, interpretability, and acceptability to create a machine. In addition, to avoid a risk of harm to human life from an autonomous system, researchers need to formulate a theory of how to allocate responsibility for outcomes produced by functionally autonomous AI technologies, like autonomous vehicles and autonomous weapons.

In the connected world, WI-related research and development will eventually enter into the "one" era. It means everything in WI would like to work at one mode, being integrated into and contributing to one social-cyber-physical space, and the one produces three ways of the ABC, and the three motivate the unlimited potential with the future of everything. A detailed case is introduced through integrating the *HIGH5* scheme into the road of constructing the future intelligence society, which encourages us to embrace each other as ones, not singles. WI 3.0 should take the responsibility to better benefit humankind, making it possible to accurately promote the creation of sci-tech innovation, and applying smoothly them with the considerations such as trustworthy, reliability, and explainability.

Appendix 1: evolution of WI tracks and topics

Evolution of tracks and topics from WI 1.0 to WI 3.0 is given from the expert's viewpoints.

Tracks and topics of WI 1.0

What are issues and research topics on WI 1.0? In order to study advanced Web technology systematically, and develop advanced Web-based intelligent information systems, we give an overview of WI related topics and list several major subtopics in each topic below.

Intelligent Web-Based Business: Business Intelligence; Customer Relationship Management; Electronic Commerce and Electronic Business; Measuring and Analyzing Web Merchandising; Price Dynamics and Pricing Algorithms; Targeted Marketing; Web-Based EDI; Web Marketing; Web Publishing; Web Services.

Knowledge Networks and Management: Electronic Library; Information and Knowledge Markets; Network Community Formation and Support; Ontology Engineering; Semantic Web; Visualization of Information and Knowledge; Web-Based Decision Support; Web Regularities and Models.

Ubiquitous Computing and Social Intelligence: Competitive Dynamics of Web Sites; Computational Societies and Markets; Dynamics of Information Sources; Reputation Mechanisms; Social Networks; Theories of the Small-World Web; Ubiquitous Learning Systems; Ubiquitous Web Access; Web-Based Cooperative Work; Web Security, Integrity, Privacy and Trust; Wireless Web Intelligence.

Intelligent Human-Web Interaction: Adaptive Web Interfaces; Multimodal Data Processing; Multimedia Representation; Science and Art of Web Design.

Web Information Management: Data Models for the Web; Internet and Web-Based Data Management; Multi-Dimensional Web Databases and OLAP; Multimedia Information Management; Object-Oriented Web Information Management; Personalized Information Management; Semistructured Data Management; Use and Management of Metadata; Web-Based Distributed Information Systems.

Web Information Retrieval: Automatic Cataloging and Indexing; Conceptual Information Extraction; Multimodal Information Retrieval; Multilinguistic Information Retrieval; Multimedia Retrieval; Ontology-Based Information Retrieval; Information Retrieval Support Systems.

Web Agents: Conversational Systems; E-mail Filtering and Automatic Handling; Global Information Foraging; Information Filtering; Navigation Guides; Recommender Systems; Resource Intermediary and Coordination Mechanisms; Remembrance Agents; Semantic Web Agents.

Web Mining and Farming: Data Mining and Knowledge Discovery for WI; Learning User Profiles; Multimedia Data Mining; Text Mining; Web-Based Ontology Learning; Web-Based Reverse Engineering; Web Farming; Text Categorization; Web-Content Mining; Web-Log Mining; Web-Structure Mining; Web Warehousing.

Emerging Web Technology and Infrastructure: Grid Computing; New Web Information Description and Query Languages; Peer-to-Peer Computing; Problem Solver Markup Language (PSML); Soft Computing (including neural networks, fuzzy logic, evolutionary computation, rough sets, and granular computing) and Uncertainty Management for WI; Web

Document Prefetching; Web Inference Engine; Web Intelligence Development Tools; Web Protocols; Wisdom Web.

Tracks and topics of WI 2.0

What are issues and research topics on WI 2.0? In order to study advanced Web technology systematically, and develop advanced Web-based intelligent information systems, we give an overview of WI related topics and list several major subtopics in each topic below.

Web of People: Crowdsourcing and Social Data Mining; Human Centric Computing; Information Diffusion; Knowledge Community Support; Modelling Crowdsourcing; Opinion Mining; People Oriented Applications and Services; Recommendation Engines; Sentiment Analysis; Situational Awareness Social Network Analysis; Social Groups and Dynamics; Social Media and Dynamics; Social Networks Analytics; User and Behavioural Modelling.

Web of Data: Algorithms and Knowledge Management; Autonomy-Oriented Computing (AOC); Big Data Analytics; Big Data and Human Brain Complex Systems; Cognitive Models; Computational Models; Data Driven Services and Applications; Data Integration and Data Provenance; Data Science and Machine Learning; Graph Isomorphism; Graph Theory; Information Search and Retrieval; Knowledge Graph; Knowledge Graph and Semantic Networks; Linked Data Management and Analytics; Self-Organizing Networks; Semantic Networks; Sensor Networks; Web Science.

Web of Things: Complex Networks; Distributed Systems and Devices; Dynamics of Networks; Industrial Multi-domain Web; Intelligent Ubiquitous Web of Things; IoT Data Analytics; Location and Time Awareness; Open Autonomous Systems; Streaming Data Analysis; Web Infrastructures and Devices Mobile Web.

Web of Trust: Blockchain Analytics and Technologies; Fake Content and Fraud Detection; Hidden Web Analytics; Monetization Services and Applications; Trust Models for Agents; Ubiquitous Computing; Web Cryptography; Web Safety and Openness; Web-scale Security, Integrity, Privacy and Trust.

Web of Agents: Agent Networks; Autonomy Remembrance Agents; Autonomy-oriented Computing; Behavior Modelling; Distributed Problem-Solving Global Brain; Edge Computing; Individual-based Modelling Knowledge; Information Agents; Local–global Behavioral Interactions; Mechanism Design; Multi-Agent Systems; Network Autonomy Remembrance Agents; Self-adaptive Evolutionary Systems; Self-organizing Systems; Social Groups and Dynamics.

Emerging Web in Health and Smart Living: Big Data in Medicine; City Brain and Global Brain; Digital Ecosystems; Digital Epidemiology; Health Data Exchange and Sharing; Healthcare and Medical Applications and Services; Omics Research and Trends; Personalized Health Management and Analytics; Smart City Applications and Services; Time Awareness and Location Awareness Smart City; Wellbeing and Healthcare in the 5 G Era.

Tracks and topics of WI 3.0

Facing the development of future WI, we first collect and summary the expert-viewed topics distributed in different tracks from scientists within different domains such as AI, cognitive science, behavior science, as well internet and communication technology.

Web of People: Cognitive Modeling and Computing; Conversational Search and Dialog Systems; Crowdsourcing and Social Computing; Human Centric Computing and Services; Human Creativity and Decision-making Support; Human-level Collective Intelligence; Human-machine Co-intelligence in the Connected World; Information Diffusion Modeling and Analysis; Opinion Mining and Sentiment Analysis; Recommendation Systems; Situation and Personality Awareness; Social Media and Social Networks; User and Behavioral Modelling; Wisdom Services.

Web of Data: Artificial Intelligence Generated Content (AIGC); Big Data Analytics and Deep Learning; Big Data and Human Brain Complex Systems; Cognitive Models and Computational Models; Data Driven Services and Applications; Data Integration and Data Provenance; Data-Knowledge-Wisdom Hierarchy; Data Science and Machine Learning; Few-shot Learning and Transfer Learning; Graph Isomorphism and Graph Theory; Information Search and Retrieval; Knowledge Graph and Semantic Networks; Linked Data Management and Analytics; Multimodal Data Fusion; Pre-trained Language Models and Applications; Representation Learning.

Web of Things: Distributed Systems and Devices; Dynamics of Networks; Industrial Multi-domain Web; Intelligent Ubiquitous Web of Things; IoT Data Analytics; Location and Time Awareness; Open Autonomous Systems; Sensor Networks; Streaming Data Analysis; Web Infrastructures and Devices Mobile Web; Wisdom Web of Things.

Web of Trust: Blockchain Analytics and Technologies; De-Platforming and No-platforming; Decentralization of Internet; Fake Content and Fraud Detection; Hidden Web Analytics; Monetization Services and Applications; Trust Models for Agents; Ubiquitous Computing; Web Cryptography; Web Safety and Openness; Web-scale Security, Integrity, Privacy and Trust.

Web of Agents: Agent Networks and Multi-Agent Systems; Autonomy Remembrance Agents; Autonomy-oriented Computing; Behavior Modelling and Individual-based Modelling; Chatbot and Intelligent Agent; Computational Social Science; Deep Reinforcement Learning; Distributed Problem-Solving and Reasoning; Edge Computing and Cloud Computing; Local–global Behavioral Interactions; Mechanism Design; Network Autonomy Remembrance Agents; Self-adaptive and Self-organizing Evolutionary Systems; Social-cyber-physical Systems; Symbols-Meaning-Value Space.

Web in Industry, Society, Health and Smart Living, and the Web of Everything: AIGC in Industry, Finance, Culture, Tourism, Education and Healthcare; Data Brain, City Brain and Global Brain; Data-driven Service Industry; Data-driven Innovative Service-oriented Society; Digital Ecosystems and Digital Epidemiology; Digital Transformation and Digital Twin; Generative AI and the Web of Everything; Human–machine Symbiosis in a Connected World; Web3, Metaverse and Smart Living; Wellbeing and Healthcare in the 5 G Era.

FAccT and AIGC: Fairness, Accountability, and Transparency; Generative AI; Explainability and Interpretability; Responsible AI; Metric and Evaluation; Applications and Use Cases; Impact on Society; Robustness and Security.

Appendix 2: topic trend analysis method

For an original topic corpus $\{o.\}$ with the temporal characteristic, it can be divided to different subsets $\{\mathbb{OT}_{S_1}, \ldots, \mathbb{OT}_{S_i}, \ldots, \mathbb{OT}_{S_j}, \ldots\} = \mathbb{OTS}$ depending on predefined parameters of a starting point $\mathcal{X}$, a end point $\mathcal{Z}$, a step size $\mathcal{Y}$, and a time window $\mathcal{W}$ as initialization firstly, corresponding to original records. Each subset $\mathbb{OT}_{S_i}$ includes the records $\{o_{\mathcal{X}+i\mathcal{Y}}, \ldots, o_{\mathcal{X}+i\mathcal{Y}+\mathcal{W}} | i \in \mathbb{Z}_0^+, \mathcal{X} + i\mathcal{Y} + \mathcal{W} \leq \mathcal{Z}\}$, where the selected records from $\{o.\}$ are related to different time ranges of $[\mathcal{X} + i\mathcal{Y}, \mathcal{X} + i\mathcal{Y} + \mathcal{W}]$. Meanwhile, the *CoGC* in Fu3ML can be used to generate new content $\{r_\bullet\}$ on the basis of prompts created by the *EnRC* in Fu3ML and the manual methods from individuals. Herein, the prompt is the keywords from each subset $\mathbb{OT}_{S_i}$, generating the corresponding extended subset $\mathbb{PT}_{S_i} \in \mathbb{PTS} = \{\mathbb{PT}_{S_1}, \ldots, \mathbb{PT}_{S_i}, \ldots, \mathbb{PT}_{S_j}, \ldots\}$. Such generative contents can be regarded as an extension of investigative topics during a phase, which can be regarded as the next-phase prediction corresponding to the future investigative trends at the same time.

Furthermore, the embeddings of original contents and their predictions are calculated by text encoding methods for $\mathbb{OT}_{S_i}$ and $\mathbb{PT}_{S_i}$, generating $\mathbb{EOT}_{S_i} = TextEncoder(\mathbb{OT}_{S_i})$ and $\mathbb{EPT}_{S_i} = TextEncoder(\mathbb{PT}_{S_i})$, respectively. Next, we integrate three strategies to measure the topic switching degree $\varepsilon = \Delta_1\alpha + \Delta_2\beta + \Delta_3\gamma$ through measuring the semantic similarity $Func\_SSM(\cdot,\cdot)$ of identified topics distributed in different parametric spaces, where the sum of Δ_1, Δ_2 and Δ_3 is 1. Such three strategies can be formulated by the following equations:

$$\begin{cases} \alpha_i = Func\_SSM(\mathbb{EOT}_{S_i}, \mathbb{EOT}_{S_{i+1}}), \\ \beta_i = Func\_SSM(\mathbb{EPT}_{S_i}, \mathbb{EPT}_{S_{i+1}}), \\ \gamma_i = Func\_SSM(\mathbb{EPT}_{S_i}, \mathbb{EOT}_{S_{i+1}}). \end{cases} \tag{B.1}$$

where the different parameters and operations for topic learning within and among different time windows can impact the size of topic shifting degrees $\varepsilon_i = \Delta_1\alpha_i + \Delta_2\beta_i + \Delta_3\gamma_i$ with the scales of $0 \leqslant \varepsilon_i, \alpha_i, \beta_i, \gamma_i \leqslant 1$, and then help us understand topic switching points.

It can be found that the smaller a shifting degree ε_i is, the bigger the topic switching is. Such processes iterate from the start time-points to end time-points with predefined parameters. Then, the calculated ε_i values are averaged to $\bar{\varepsilon}$ under different conditions corresponding to each predefined time window. Finally, we can identify the balanced quantitative value from all $\bar{\varepsilon}$ values like Moore's Law to measure the time switching scale of investigative topics, through the comparison of all time windows. Detailed workflows are given in Algorithm 1 of Appendix 2.

Input: An original topic corpus $\{o_\bullet\}$ built by the $EnRC$ in Fu3ML; A starting point $\mathcal{X}$; A end point $\mathcal{Z}$;
A step size $\mathcal{Y}$; A time window $\mathcal{W}$;
The weights of $\varDelta_1$, $\varDelta_2$, and $\varDelta_3$;
The size of inspired contents $\varUpsilon$;
The size of extended contents $\varPhi$;

Output: The set of topic switching degrees $\varepsilon = \{\varepsilon_i\}$ corresponding to the current parameters; The topic switching degree $\bar{\varepsilon}$ calculated by mean of the ε at current time scale.

Let $\varepsilon = \varnothing$; $\alpha = \varnothing$; $\beta = \varnothing$; $\gamma = \varnothing$; $i = 0$;
Let $\mathbb{EOTS} = \varnothing$; ▷ The set of embeddings on original topics
Let $\mathbb{EPTS} = \varnothing$; ▷ The set of embeddings on extended topics
if $\mathcal{X} + (i+1)\mathcal{Y} + \mathcal{W} - 1 \leq \mathcal{Z}$ **then**
  while $\mathcal{X} + i\mathcal{Y} + \mathcal{W} - 1 \leq \mathcal{Z}$ **do**
    select $\mathbb{OT}_{S_i}$ with the size of $\varUpsilon$ from $\{o_\bullet\}$ in $[\mathcal{X} + i\mathcal{Y}, \mathcal{X} + i\mathcal{Y} + \mathcal{W})$, and add it into $\mathbb{OTS}$;
    extend $\mathbb{OT}_{S_i}$ to $\mathbb{PT}_{S_i}$ with the size of $\varPhi$, and add it into $\mathbb{PTS}$;
    compute the embeddings of inspired contents, $\mathbb{EOT}_{S_i} = EmLC(\mathbb{OT}_{S_i})$, and add it into $\mathbb{EOTS}$;
    compute the embeddings of extended contents, $\mathbb{EPT}_{S_i} = EmLC(\mathbb{PT}_{S_i})$, and add it into $\mathbb{EPTS}$;
    $i = i + 1$;
  end while
  K = the length of $\mathbb{EOTS}$;
  for $i = 0$ **to** K **do**
    $\alpha_i = Func\_SSM(\mathbb{EOT}_{S_i}, \mathbb{EOT}_{S_{i+1}})$;
    $\beta_i = Func\_SSM(\mathbb{EPT}_{S_i}, \mathbb{EPT}_{S_{i+1}})$;
    $\gamma_i = Func\_SSM(\mathbb{EPT}_{S_i}, \mathbb{EOT}_{S_{i+1}})$;
    $\varepsilon_i = \varDelta_1\alpha_i + \varDelta_2\beta_i + \varDelta_3\gamma_i$;
    add ε_i into ε;
  end for
  compute $\bar{\varepsilon}$ from ε;
end if
Save the aligned $\langle \mathbb{OTS}, \mathbb{PTS}, \mathbb{EOTS}, \mathbb{EPTS} \rangle$;
Return the ε and $\bar{\varepsilon}$.

Algorithm 1 Topic Switching Point-oriented Detection and Prediction (TSPDP)

Furthermore, the future trends of WI 3.0 can be predicted by using generative AI models, combining with TSPDP and FDETSD defined in Algorithm 2 of Appendix 2. Overall of pipeline is given as follows:

- calculating the averaged topic-oriented time switching degrees $[\bar{\varepsilon}\ in\ \mathcal{W}_i]$ learned by Algorithm 1 of Appendix 2 under various $\mathcal{W}$ corresponding to different parameters $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$;
- selecting the parameters $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, $\mathcal{W}$ corresponding to a $\bar{\varepsilon}$, considering the balance of time windows and jumping scales;
- predicting the future research trends via Algorithm 2 of Appendix 2 with the selected parameters, combining with the *CoGC* in Fu3ML.

Input: The original topic corpus $\{o_\bullet\}$ built by the $EnRC$ in Fu3ML; The scale of extended steps $\nu_O \in \mathbb{Z}^+$;
▷ for the trend prediction from the temporal view
The inspired topic corpus $\{r_\bullet\}$; The scale of extended tracks $\nu_R \in \mathbb{Z}^+$;
▷ for the trend prediction from the scope view
The $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ and $\mathcal{W}$ corresponding to a $\bar{\varepsilon}$ via balancing multiple $\mathcal{Y}$ and $\mathcal{W}$ results learned by Algorithm 1 of Appendix B;
The size of inspired topics $\varUpsilon$;
The size of extended contents $\varPhi$;
Output: The set of extended topics at each step $\mathbb{ETS} = \{ETS[i] \mid i = 1, 2, \ldots, \nu_O\}$;
The set of extended topics at each track $\mathbb{ETT} = \{ETT[i] \mid i = 1, 2, \ldots, \nu_R\}$.

/ *** **Future Directions Extended in Temporal Domains** *** /
Let $\mathbb{ETS} = \varnothing$;
Identify the topic clusters from $\{o_\bullet\}$ to $\{o_i \mid i = 1, 2, \ldots\}$ based on $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ and $\mathcal{W}$;
Count the length of $\{o_i\}$, K;
Generate increasing K numbers, $\{\partial_i | i = 1, \ldots, K\}$, where their sum is 1;
Calculate $\mathrm{Int}(\partial_i * \varUpsilon)$;
for $i = 1$ **to** ν_O **do**
    extract the current topics $\mathbb{CTT}$ from $\{o_i\}$ with the size of $\mathrm{Int}(\partial_i * \varUpsilon)$;
    generate the future topics $ETS[i]$ with the size of $\varPhi$ based on $\mathbb{CTT}$ via the $CoGC$ in Fu3ML;
    add $ETS[i]$ into $\mathbb{ETS}$ and $\{o_i\}$;
    delete the first cluster from $\{o_i\}$;
end for
Return the $\mathbb{ETS}$.

/ *** **Future Directions Extended in Spatial Domains** *** /
Let $\mathbb{ETT} = \varnothing$;
Identify the topic clusters from $\{r_\bullet\}$ to $\{r_i \mid i = 1, \ldots, \nu_R\}$;
for $i = 1$ **to** ν_R **do**
    extract the current topics $\mathbb{CTT}$ from $\{r_i\}$ with the size of $\varUpsilon$;
    generate the future topics $ETT[i]$ with the size of $\varPhi$ based on $\mathbb{CTT}$ via the $CoGC$ in Fu3ML;
    add $ETT[i]$ into $\mathbb{ETT}$;
end for
Return the $\mathbb{ETT}$.

Algorithm 2 Future Directions Extended by Topics within Temporal and Scope Domains (FDETSD)

Appendix 3: experimental details

The corpus is constructed by the flow diagram obtained from PRISMA guidelines Page (2021), as shown in Fig. 3 with the number of selected snippets in abstracts at different years. In this paper, the snippets of abstracts are extracted from the studies containing the keyword "Web Intelligence", where the searching engine is based on Google Scholar for relevant publications, including journal, conference, and patents. All records of identified publications were screened by human beings for eligibility. Records had to meet the following inclusion criteria: (1) records had to include titles that indicate the topics of studies; and (2) records had to include snippets of abstracts, such as the first sentence of abstract that indicates the background of studies, these two aspects of which can give a global view for highlighted trends during different time periods at different tracks. Records were excluded based on the following criteria: (1) studies written in non English and error code; (2) failure to obtain snippets of abstracts simultaneously; and (3) the duplicate records. In the initial phase, a total of 72427 records are identified from Google Scholar via the GoogleScholarly interface. Depending on the standard of maximum relevance, top 1000 records are extracted for each year, including any language and any type of papers. Then, we remove non-English, duplicate, and empty records, the sizes of eligible paper are given in Table 1 of Appendix 3. Finally, a total of 17902 records (including the papers' snippets in abstracts) are selected to the next analytical phase.

During the analysis processes in practice, the screened snippets (usually including the first two sentences of abstracts) are used to decode and predict trends with topics for WI, which contain sufficient text with background and direction information. On the basis of the corpus, several experiments are performed by combining the following pre-trained foundation models:

Table 1 Overview of the publication scale on Web Intelligence from 2000 to 2023

YEAR	Proc				
	Proc[1]	Proc[2]	Proc[3]	Proc[4]	Proc[5]
2000 [FULL]	44	42	42	35	31
2001 [...]	172	170	169	153	140
2002 [...]	173	170	164	144	130
2003 [...]	595	590	570	484	449
2004 [...]	933	926	905	826	770
2005 [...]	1270	999	960	874	813
2006 [...]	1590	998	957	866	838
2007 [...]	1880	1000	957	879	845
2008 [...]	2450	1000	981	904	890
2009 [...]	2760	1000	977	919	899
2010 [...]	3560	999	986	900	884
2011 [...]	3510	999	976	896	871
2012 [...]	3820	999	973	866	848
2013 [...]	4010	998	968	897	883
2014 [...]	4430	1000	966	895	876
2015 [...]	4560	1000	963	894	876
2016 [...]	4320	999	961	914	890
2017 [...]	4390	1000	962	920	888
2018 [...]	4600	998	961	939	830
2019 [...]	4680	995	948	923	799
2020 [...]	4900	1000	968	951	778
2021 [...]	5150	996	976	966	851
2022 [...]	5100	997	968	963	892
2023 [FULL]	3530	998	958	954	931
Total	72427	20873	20216	18962	17902

[1]The number of original records returned from the Google Scholar database through the search keyword "web intelligence"

[2]The number of extracted records via interfaces with the maximum of 1000 in each year, sorted by relevance

[3]The number of screened records that only includes English

[4]The number of screened records that are not empty

[5]The number of screened records in which duplication has been removed

- To obtain the entity information, the BERT base model (cased), that is the pre-trained "bert-base-cased" model in the Python environment, is trained to tag the properties of words in sentences via a entity annotated corpus from "bookcorpus" and "wikipedia" in a named entity recognition task, where the words with the none property are extracted to represent points of interest in a text;
- To obtain the significant entities, the ConSERT framework is firstly used to compute embeddings of identified entities by combining the pre-trained "text2vec-base-multilingual" model via the "text2vec" library, and then the "KMeans" method is executed to decode the centered keywords from different clusters in the "scikit-learn" environment of Python, where the selected centered keywords are demonstrated as highlights of a text;
- To predict the future directions, the GPT interface is called by combining the pre-trained "text-davinci-003" model with the prompts of from the OpenAI team, where **k** is 10 that controls the extended scale for each entity from **foc_class**, the **foc_class** is a set organized by the manually and/or automatically obtained entities with the preset sampled size, with respect to different time windows and tracks.

$$
\begin{aligned}
&\text{``}\\
&\textbf{prompt} =\\
&\quad \textit{These are} < \textbf{foc\_class} > \textit{entity names.}\\
&\quad \textit{Generate } \textbf{k} \textit{ new} < \textbf{foc\_class} > \textit{entity names.}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{''}
\end{aligned}
$$

Through combining the above three sets of parameters, the entities are reviewed and extended from the time evolution and domain views. In particular, from the time view, the original corpus is split up into different subsets with the time window property of yearly basis from different time intervals of [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. Then the original entities are respectively extracted and then are extended to obtain new entities. Therefore, the topic switching degrees are detected by Algorithm 1 with respect to the different time windows and jumping scales, using original entities and extended entities. The jumping scale is from [1, 2, 3, 4, 5], which represents the gap size of time windows between the original entities and the extended entities. Through balancing the size of the time windows and the jumping scales, the topic switching point is detected to demonstrate the updated and iterated scales of WI research. Furthermore, the limited entities are sampled within a fixed time window based on the selected topic switching point, where the size of sampled entities in each time window is increased from past to recent through running the *numpy.linspace*() function, the sum of which is 300. With the generation of new entities, the entities in the left time window are deleted at the earliest, and the new entities are appended to the right side at the same time, where the processes are iterated with eight times corresponding to $[step-1, \ldots, step-8]$, respectively. The extended entities followed by different time steps can inspire our next-stage works. Meanwhile, from the domain view, the future trends of WI are predicted under the different tracks in the WI 3.0 era corresponding to $[track-1, \ldots, track-7]$, respectively. Herein, all topics are organized as prompts under different tracks, running with the GPT interface. To observe the key points and strengthen the flexibility of visualization, the entities' embeddings are computed and then are clustered to observe the highlighted research activities within different time windows and tracks. Herein, the sizes of clusters are designed to "six" during the reviewing processes and "fifteen" during the predicting processes.

Appendix 4: experimental results

From Figs. 8 to 14, the odd rows represent the original entities with the sign of *O* and the even rows represent their corresponding extended entities with the sign of *P*, with respect to Web Intelligence-related studies from 2000 to 2023. The comparison of the original and extended entities is illustrated in the following figures:

- In Fig. 8 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized annually;
- In Fig. 9 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized every 2 years, without the overlapping windows;
- In Fig. 10 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized every 3 years;
- In Fig. 11 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized every 4 years from one to four rows, and every 5 years from five to eight rows;
- In Fig. 12 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized every 6 years from one to four rows, and every 7 years from five to eight rows;
- In Fig. 13 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized every 8 years from one to four rows, and every 9 years from five to eight rows;
- In Fig. 14 of Appendix 4, the top-6 topics corresponding to the center of different clusters are visualized with the time window of 10 years.

With the increasing of time windows, the occasional nature of original entities that are sampled becomes higher, which further increases the diversity of extended entities. For example, corresponding to two sets of similar original entities, their extended results exhibit relative higher similarity in smaller time windows such as the three years' window in Fig. 10 than that of bigger time windows such as the ten years' window in Fig. 14.

In Fig. 15 of Appendix 4, the trend shift degrees are given at the different time windows from 1 to 10 years and the different jumping scales from 0 to 4 years. Whether the time window or the jumping scale, the trend shift degrees become smaller and smaller as the values grows, which indicate the trend shift becomes bigger and bigger. In particular, three critical turning points appear in the time windows of 3 years, 4 years, and 5 years. Taken together, adapted to both settings of time windows and jumping scales, it can be found that the significant iteration and updating in the WI field can be achieved every 4 years.

In Fig. 16 of Appendix 4, the fifteen core topics are created with the sequencing attribute from clustered results that are obtained by the embedded and extended entities from Step-1 to Step-8, along with the moving of time windows from early to recent.

In Fig. 17 of Appendix 4, the fifteen core topics are created with the scope attribute from clustered results that are obtained by the embedded and extended entities corresponding to different tracks from track-1 to track-7, inspired by the experts' opinions, respectively.

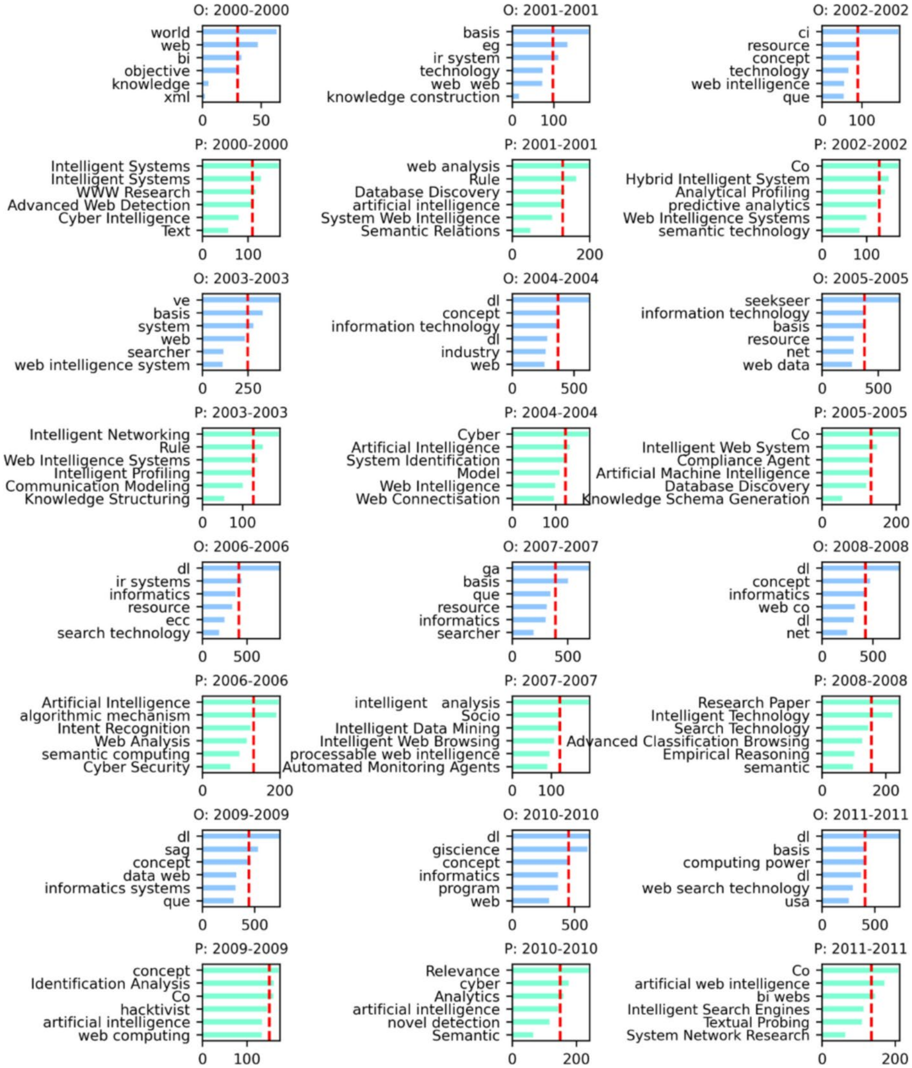

Fig. 8 The centric entities are identified annually from 2000 to 2023. The original entities are at the odd rows and their corresponding extended entities are at the even rows

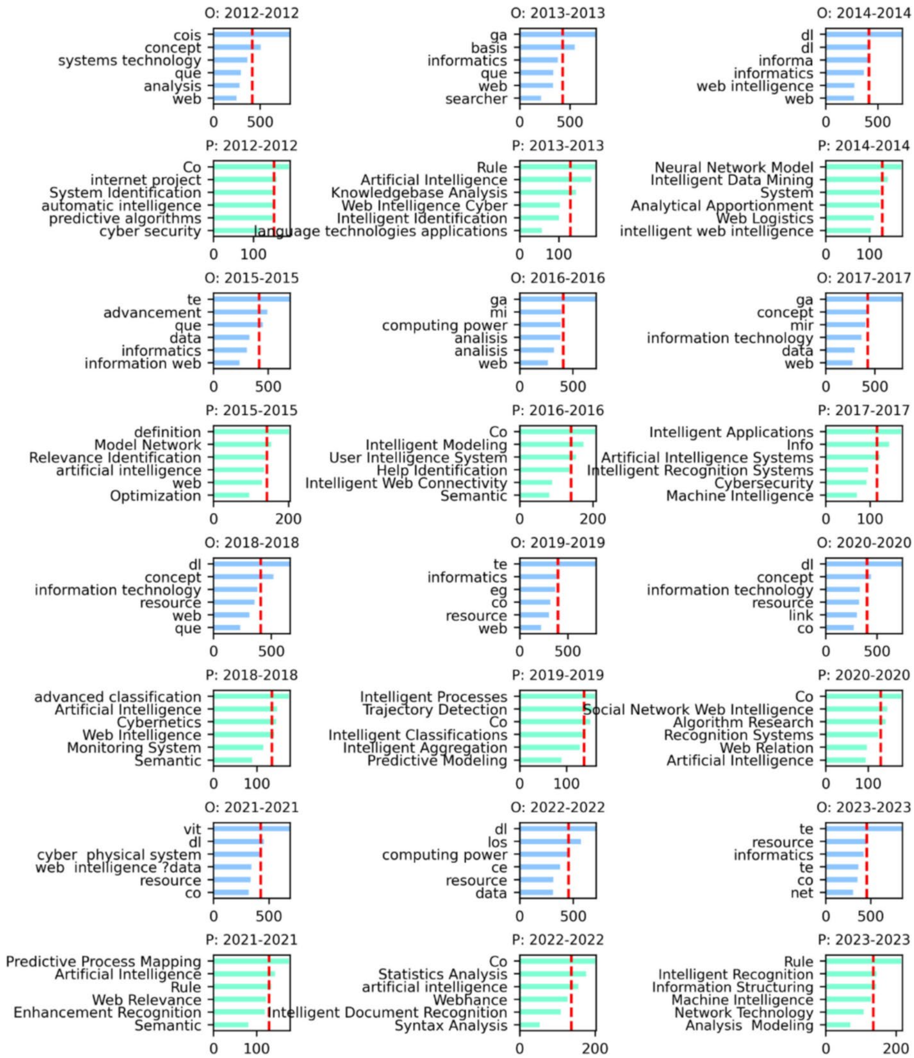

Figure 8 (continued)

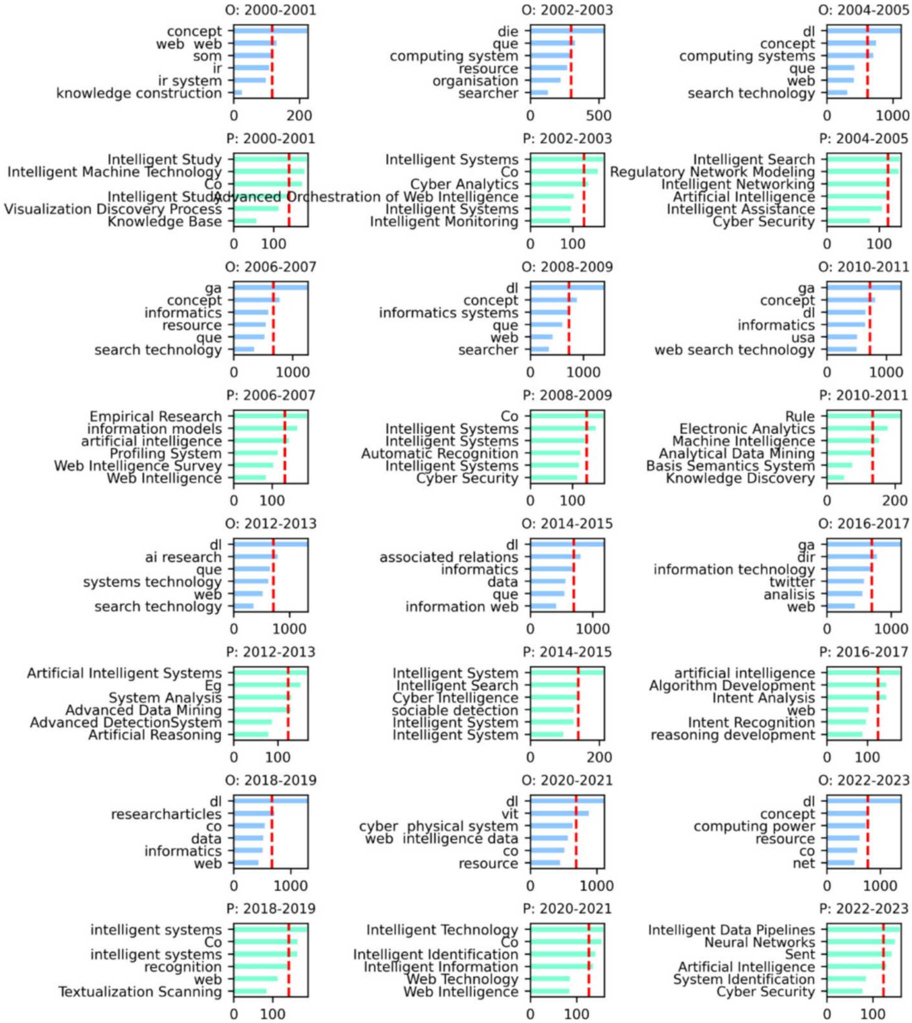

Fig. 9 The centric entities are identified every 2 years from 2000 to 2023. The original entities are at the odd rows and their corresponding extended entities are at the even rows

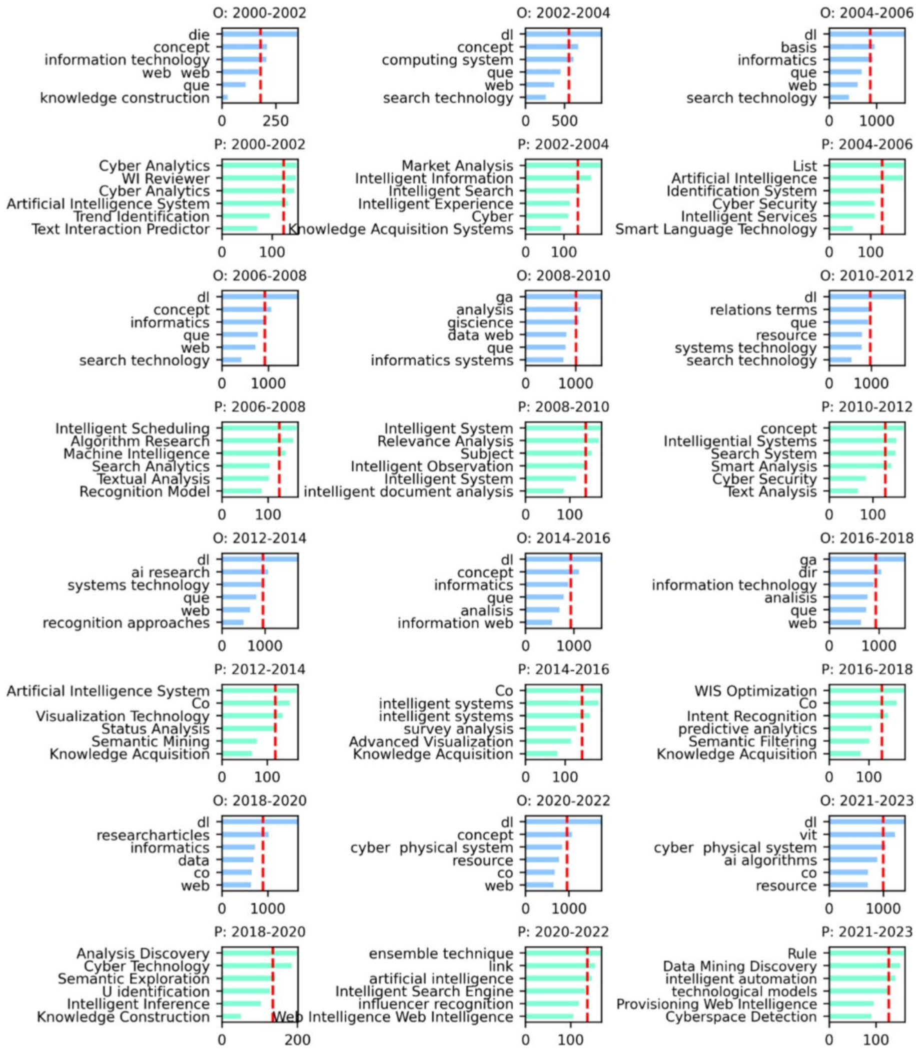

Fig. 10 The centric entities are identified every 3 years from 2000 to 2023. The original entities are at the odd rows and their corresponding extended entities are at the even rows

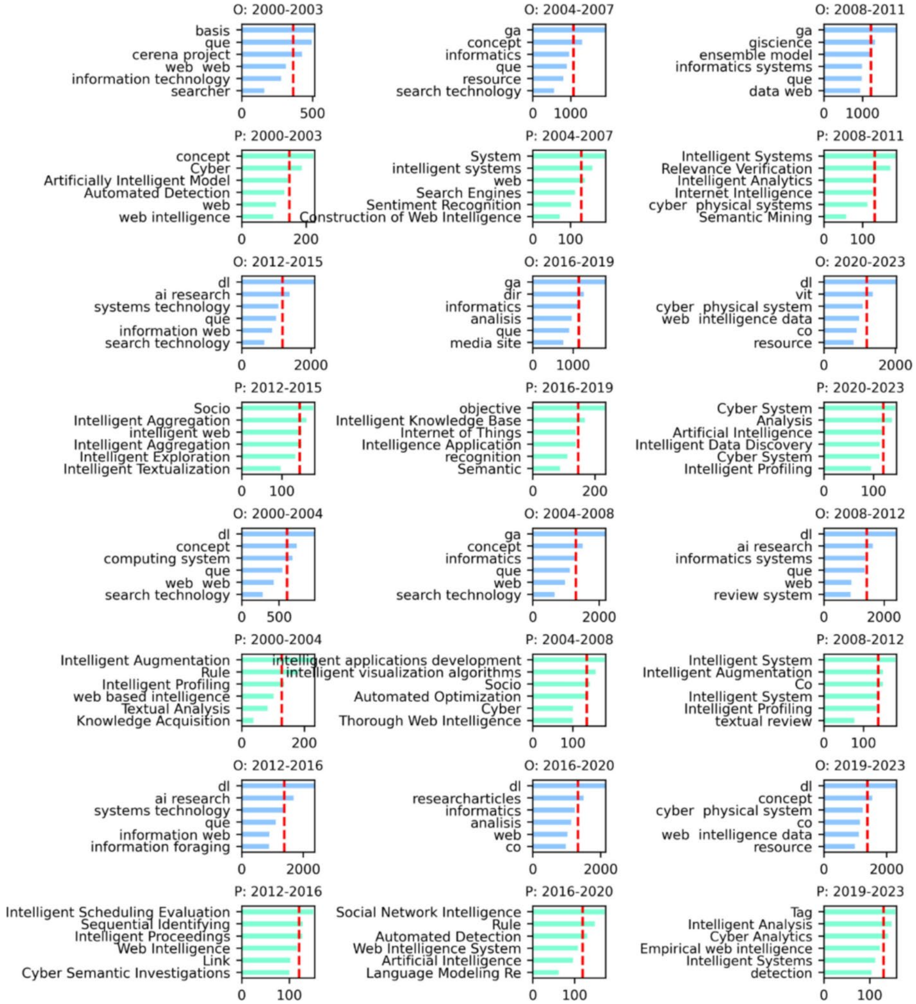

Fig. 11 The centric entities are identified every 4 years from one to four rows and every 5 years from five to eight rows. The original entities are at the odd rows and their corresponding extended entities are at the even rows

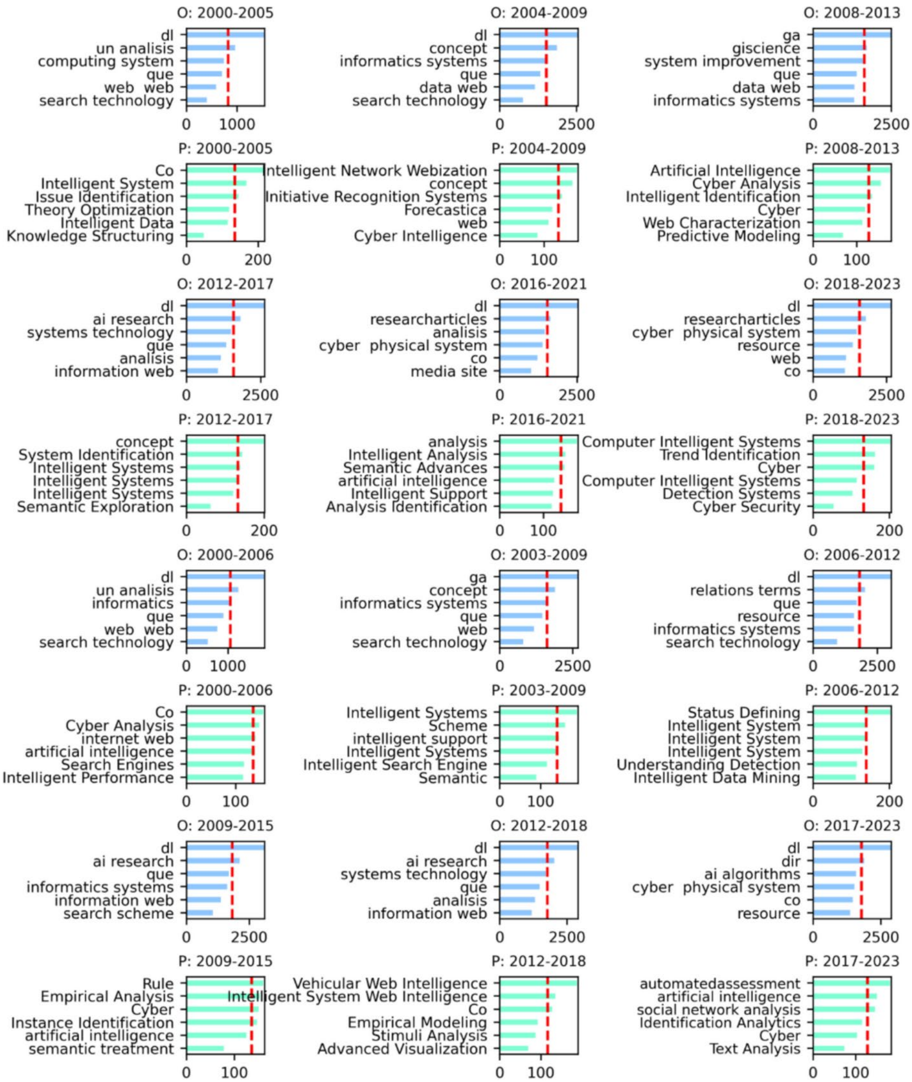

Fig. 12 The centric entities are identified every 6 years from one to four rows and every 7 years from five to eight rows. The original entities are at the odd rows and their corresponding extended entities are at the even rows

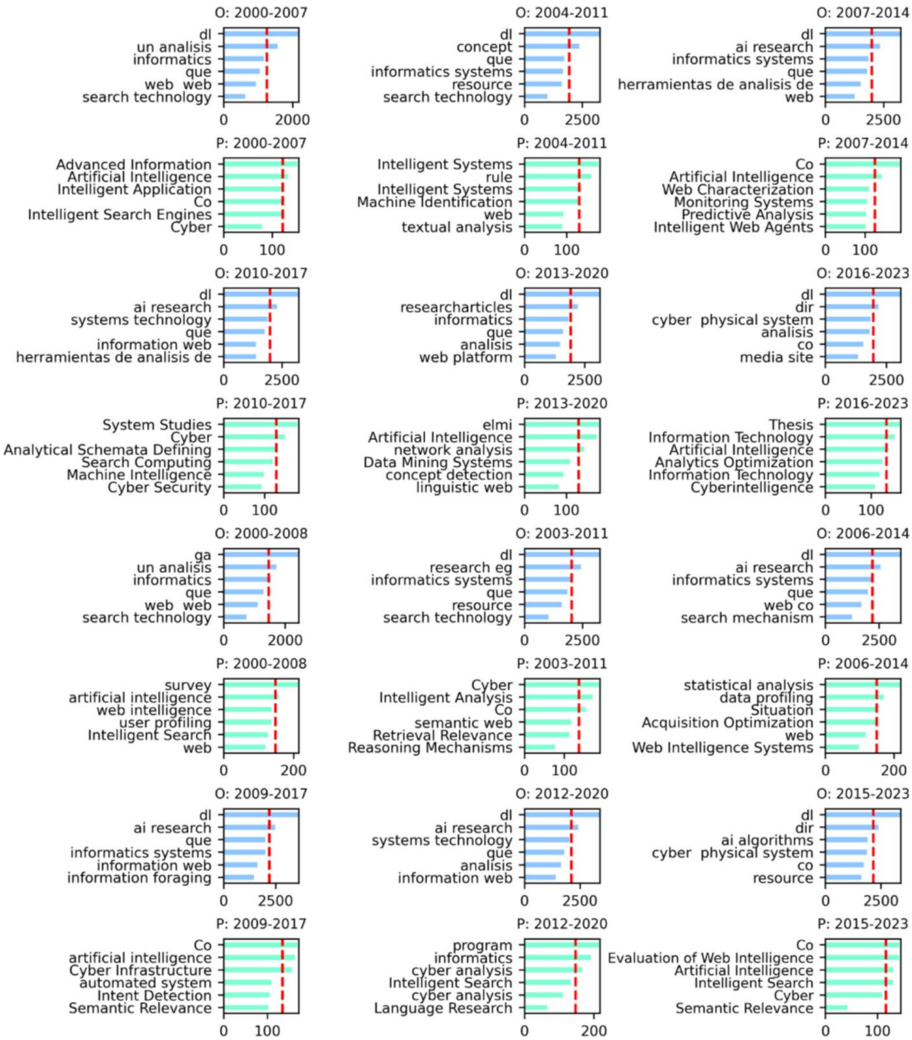

Fig. 13 The centric entities are identified every 8 years from one to four rows and every 9 years from five to eight rows. The original entities are at the odd rows and their corresponding extended entities are at the even rows

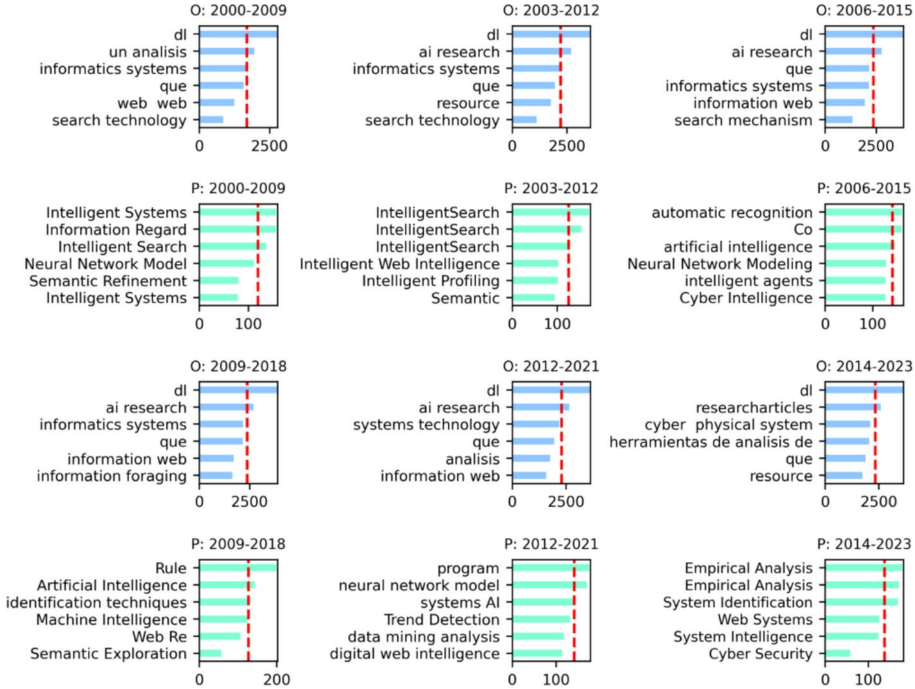

Fig. 14 The centric entities are identified every 10 years from 2000 to 2023. The original entities are at the odd rows and their corresponding extended entities are at the even rows

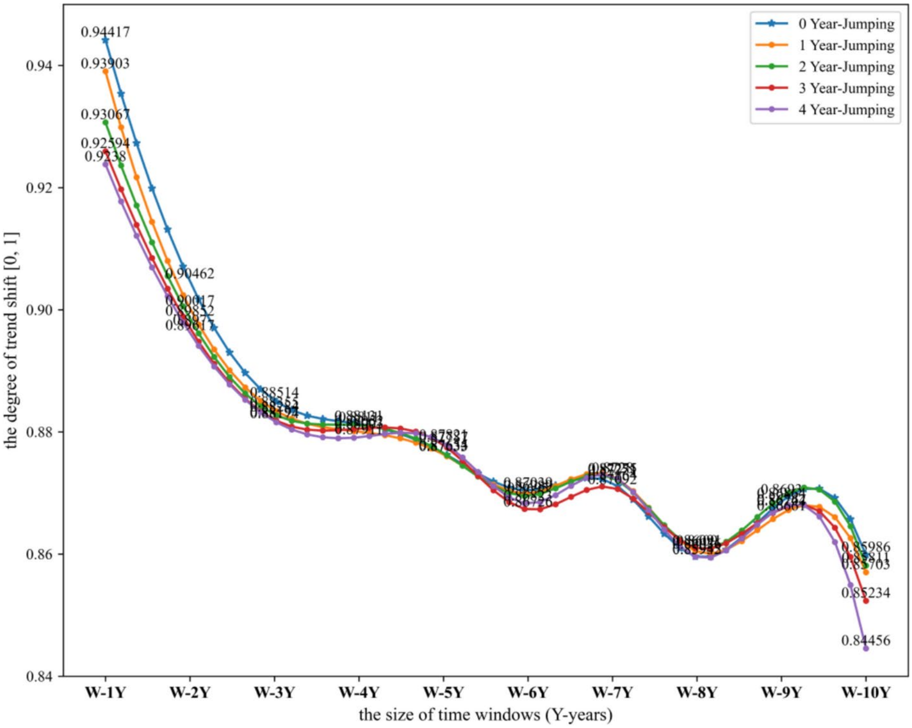

Fig. 15 Review of trend shift degrees in WI from 2000 to 2023, with respect to the different time windows from 1 to 10 years and the jumping scales from 0 to 4 years

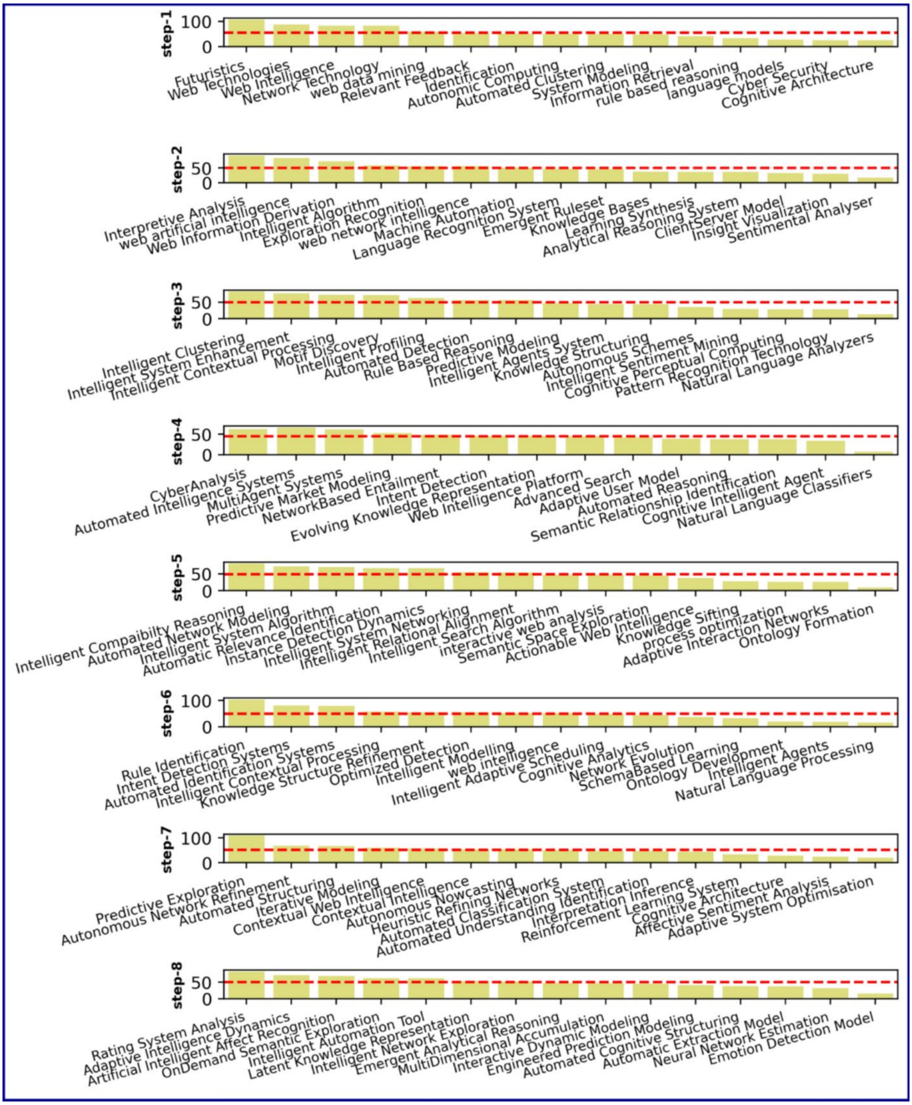

Fig. 16 Extension of topics followed by the increased time scales in WI from step-1 to step-8, where the time window is 4 years selected from Fig. 15 without overlapping

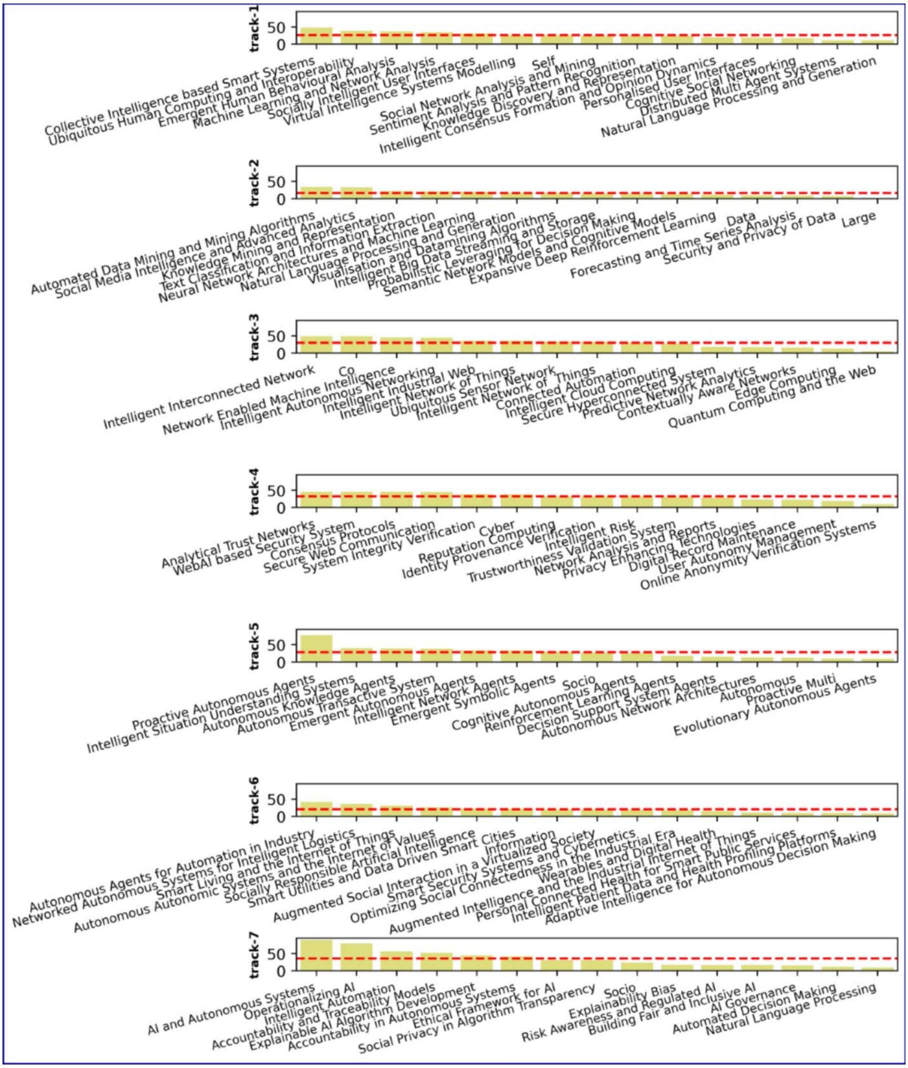

Fig. 17 Extension of topics corresponding to the seven tracks in WI from track-1 to track-7. track-1: Web of People; track-2: Web of Data; track-3: Web of Things; track-4: Web of trust; track-5: Web of Agents; track-6: Web in Industry, Society, Health and Smart Living, and the Web of Everything; and track-7: Fairness, Accountability, Transparency, and AI Generated Content

Acknowledgements This research is supported by the research grants from JSPS Grants-in-Aid for Scientific Research of Japan (19 K12123); Natural Sciences and Engineering Research Council (NSERC) of Canada and York Research Chairs (YRC) program. We thank to the renowned researchers who have supported and contributed to the WI field, including E. A. Feigenbaum, L. A. Zadeh, J. McCarthy, T. M. Mitchell, S. Ohsuga, K. Friston, R. M. Karp, Y. Anzai, J. Hopcroft, A. C.-C. Yao, J. Sifakis, B. Lampson, L. Valiant, R. Reddy, and F. van Harmelen.

Author contributions H. Kuai wrote the original draft and conducted the data curation, formal analysis and visualization. H. Kuai, Y. Yao, J. Liu, and N. Zhong made contributions to the conception and methodology. H. Kuai, J. Huang, X. Tao, G. Pasi, and N. Zhong identified the potential issues and directions from the results, and contributed to formatting of the work. All the authors reviewed the manuscript.

Data availability The dataset for this research is available upon obtaining an approval from the corresponding author.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdul-Ghafour S, Ghodous P, Shariat B, Perna E, Li Y, Raghavan VV (2007) A common design-features ontology for product data semantics interoperability. In Li Y, Raghavan VV (eds) IEEE/WIC/ACM international conference on web intelligence (WI'07), 443–446 IEEE

Alam S, Dobbie G, Riddle P, Zhong N, Liu J, Yao Y (2008) Particle swarm optimization based clustering of Web usage data. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 451–454, IEEE

Angles R, Gutierrez C (2008) Survey of graph database models. ACM Comput Surv (CSUR) 40:1–39

Armbrust M et al (2010) A view of cloud computing. Commun ACM 53:50–58

Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. Comput Netw 54:2787–2805

Badue C et al (2021) Self-driving cars: a survey. Expert Syst Appl 165:113816

Balahur A, Steinberger R, Van Der Goot E, Pouliquen B, Kabadjov M, Zhong N, Liu J, Yao Y (2009) Opinion mining on newspaper quotations. In Zhong N, Liu J, Yao Y (eds), 2009 IEEE/WIC/ACM international joint conference on web intelligence and intelligent agent technology, Vol 3, 523–526, IEEE

Balazs JA, Velásquez JD (2016) Opinion mining and information fusion: a survey. Inform Fusion 27:95–110

Bannister F, Connolly R (2011) The trouble with transparency: a critical review of openness in e-government. Policy Internet 3:1–30

Bassani E, Pasi G (2022) A multi-representation re-ranking model for personalized product search. Inform Fusion 81:240–249

Baumgartner R, Frölich O, Gottlob G, Herzog M, Lehmann P, Althoff K-D, Dengel A, Bergmann R, Nick M, Roth-Berghofer T (2005) Integrating semi-structured data into business applications: a web intelligence example. In Althoff KD, Dengel A, Bergmann R, Nick M, Roth-Berghofer T (eds) Professional knowledge management: third biennial conference, WM 2005, Kaiserslautern, Germany, April 10–13, 2005, Revised Selected Papers 3, 469–482 Springer

Becker J, Brackbill D, Centola D (2017) Network dynamics of social influence in the wisdom of crowds. Proc Natl Acad Sci 114:E5070–E5076

Berkan RC, Trubatch SL (2002) Fuzzy logic and hybrid approaches to Web Intelligence gathering and information management. 2002 IEEE world congress on computational intelligence. 2002 IEEE international conference on fuzzy systems. FUZZ-IEEE'02. Proceedings (Cat. No. 02CH37291), Vol 2, 1033–1038, IEEE

Bodorik P, Jutla D, Zhong N, Liu J, Yao Y (2008) Privacy with web serivces: intelligence gathering and enforcement. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 546–549, IEEE

Borgatti SP, Mehra A, Brass DJ, Labianca G (2009) Network analysis in the social sciences. Science 323:892–895

Bradshaw JM et al (2004) Making agents acceptable to people. Intell Technol Inform Anal 361–406

Bryson JJ, Martin D, McIlraith SA, Stein LA, Zhong N, Liu J, Yao Y (2003) Agent-based composite services in DAML-S: the behavior-oriented design of an intelligent semantic web. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 37–58, Springer

Butz CJ, Hua S, Maguire RB (2006) A web-based Bayesian intelligent tutoring system for computer programming. Web Intell Agent Syst: Int J 4:77–97

Camacho D, Panizo-LLedot Ángel, Bello-Orgaz G, Gonzalez-Pardo A, Cambria E (2020) The four dimensions of social network analysis: an overview of research methods, applications, and software tools. Inform Fusion 63:88–120

Cantador I, Bellogín A, Castells P, Zhong N, Liu J, Yao Y (2008) Ontology-based personalised and context-aware recommendations of news items. In Zhong N, Liu J, Yao Y (eds) 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 1, 562–565, IEEE

Cao L (2022) Decentralized AI: edge intelligence and smart blockchain, metaverse, Web3, and DeSci. IEEE Intell Syst 37:6–19

Castellano G, Fanelli AM, Torsello MA (2008) Computational intelligence techniques for Web personalization. Web Intell Agent Syst: Int J 6:253–272

Cercone N et al (2002) From computational intelligence to web intelligence. Computer 35:72–76

Chen J et al (2014) WaaS: wisdom as a service. IEEE Intell Syst 29:40–47

Chen H, Yang B, Pei H, Liu J (2018) Next generation technology for epidemic prevention and control: data-driven contact tracking. IEEE Access 7:2633–2642

Cheng X, Gu Y, Shen J (2019) An integrated view of particularized trust in social commerce: an empirical investigation. Int J Inf Manage 45:1–12

Chen X et al (2024) Artificial intelligence and multimodal data fusion for smart healthcare: topic modeling and bibliometrics. Artif Intell Rev 57:91

Chikhi NF, Rothenburger B, Aussenac-Gilles N, Li Y, Raghavan VV, (2007) A comparison of dimensionality reduction techniques for web structure mining. In Li Y, Raghavan VV (eds), IEEE/WIC/ACM international conference on web intelligence (WI'07), 116–119, IEEE,

Deguchi A et al (2020) What is society 5.0. Society 5:1–24

Devedžić V (2004) Web intelligence and artificial intelligence in education. J Educ Technol Soc 7:29–39

Dhar V (2013) Data science and prediction. Commun ACM 56:64–73

Ding J, Petzoldt A (2017) Current state of multivariate cryptography. IEEE Security Privacy 15:28–36

Domingues MA, Jorge AM, Soares C, Leal JP, Machado P, Neves J, Santos MF, Machado JM (2007) A data warehouse for web intelligence. In Neves J, Santos MF, Machado JM (eds), Proceedings of the 13th Portuguese conference on artificial intelligence (EPIA 2007), 487–499

Elshan E, Zierau N, Engel C, Janson A, Leimeister JM (2022) Understanding the design elements affecting user acceptance of intelligent agents: past, present and future. Inf Syst Front 24:699–730

Falkowski T, Barth A, Spiliopoulou M, Li Y, Raghavan VV (2007) Dengraph: a density-based community detection algorithm. In Li Y, Raghavan VV (eds), IEEE/WIC/ACM international conference on web intelligence (WI'07), 112–115, IEEE

Faraj S, Jarvenpaa SL, Majchrzak A (2011) Knowledge collaboration in online communities. Organ Sci 22:1224–1239

Farisco M et al (2024) A method for the ethical analysis of brain-inspired AI. Artif Intell Rev 57:1–27

Faroukhi AZ, El Alaoui I, Gahi Y, Amine A (2016) Big data monetization throughout big data value chain: a comprehensive review. J Big Data 7

Feldman R (2013) Techniques and applications for sentiment analysis. Commun ACM 56:82–89

Friedewald M, Raabe O (2011) Ubiquitous computing: an overview of technology impacts. Telematics Inform 28:55–65

Fu Y, Xiang R, Liu Y, Zhang M, Ma S, Li Y, Raghavan VV (2007) Finding experts using social network analysis. In Li Y, Raghavan VV (eds), IEEE/WIC/ACM international conference on web intelligence (WI'07), 77–80 IEEE

Gao C, Su Z, Liu J, Kurths J (2019) Even central users do not always drive information diffusion. Commun ACM 62:61–67

Gao X, Wang Y, Chen X, Gao S (2021) Interface, interaction, and intelligence in generalized brain-computer interfaces. Trends Cogn Sci 25:671–684

Gu Z, Xu B, Li J (2010) Service data correlation modeling and its application in data-driven service composition. IEEE Trans Serv Comput 3:279–291

Guzdial M (2013) Human-centered computing: a new degree for Licklider's world. Commun ACM 56:32–34

Haghighi MS, Ebrahimi M, Garg S, Jolfaei A (2020) Intelligent trust-based public-key management for IoT by linking edge devices in a fog architecture. IEEE Internet Things J 8:12716–12723

Han J, Chang K-C (2002) Data mining for Web Intelligence. Computer 35:64–70

Heath T, Bizer C (2022) Linked data: evolving the web into a global data space. Springer, Cham

Hennig L, Umbrath W, Wetzker R, Zhong N, Liu J, Yao Y (2008) An ontology-based approach to text summarization. In Zhong N, Liu J, Yao Y (eds) 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 291–294, IEEE

Herschel M, Diestelkämper R, Lahmar HB (2017) A survey on provenance: What for? What form? What from? VLDB J 26:881–906

Hlinka O, Hlawatsch F, Djuric PM (2013) Distributed particle filtering in agent networks: a survey, classification, and comparison. IEEE Signal Process Mag 30:61–81

Hoeber O, Zhong N, Liu J, Yao Y (2008) Web information retrieval support systems: The future of web search. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 29–32, IEEE

Hussain A, Wenbi R, Da Silva AL, Nadher M, Mudhish M (2015) Health and emergency-care platform for the elderly and disabled people in the Smart City. J Syst Softw 110:253–263

Inui K et al, Zhong N, Liu J, Yao Y (2008) Experience mining: building a large-scale database of personal experiences and opinions from web documents. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 1, 314–321 IEEE

Jeste DV et al (2020) Beyond artificial intelligence: exploring artificial wisdom. Int Psychogeriatr 32:993–1001

Ji S, Pan S, Cambria E, Marttinen P, Yu PS (2021) A survey on knowledge graphs: representation, acquisition, and applications. IEEE Trans Neural Networks Learn Syst 33:494–514

Johnson WL, Zhong N, Liu J, Yao Y, (2003) Using agent technology to improve the quality of Web-based education. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 77–101 Springer

Jordan MI, Mitchell TM (2015) Machine Learning: trends, perspectives, and prospects. Science 349:255–260

Kalles D, Papagelis A, Zaroliagis C, Zhong N, Liu J, Yao Y (2003) Algorithmic aspects of Web Intelligent systems. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 323–344, Springer

Kambatla K, Kollias G, Kumar V, Grama A (2014) Trends in big data analytics. J Parallel Distrib Comput 74:2561–2573

Kassak O, Kompan M, Bielikova M (2018) Acquisition and modelling of short-term user behaviour on the Web: a survey. J Web Eng 17:23–70

Keyvan K, Huang JX (2023) How to approach ambiguous queries in conversational search: a survey of techniques, approaches, tools, and challenges. ACM Comput Surv 55:1–40

Knoblock CA, Szekely P (2015) Exploiting semantics for big data integration. AI Mag 36:25–38

Kong Y, Wang Y, Hajli S, Featherman M (2020) In sharing economy we trust: examining the effect of social and technical enablers on millennials' trust in sharing commerce. Comput Hum Behav 108:105993

Koukaras P, Tjortjis C, Rousidis D (2020) Social media types: introducing a data driven taxonomy. Computing 102:295–340

Kuai H, Zhong N (2020) The extensible Data-Brain model: architecture, applications and directions. J Comput Sci 46:101103

Kuai H et al (2020) THINKING-LOOP: the semantic vector driven closed-loop model for brain computing. IEEE Access 8:4273–4288

Kuai H et al (2021) Multi-source brain computing with systematic fusion for smart health. Inform Fusion 75:150–167

Kuai H, Tao X, Zhong N (2022) Web Intelligence meets Brain Informatics: towards the future of artificial intelligence in the connected world. World Wide Web 25:1223–1241

Kuai H et al (2024) Never-ending learning for explainable brain computing. Adv Sci 11:2307647

Li D, Xiao L, Han Y, Chen G, Liu K, Zhong N et al (2006) Network thinking and network intelligence. In Zhong N et al (eds), International workshop on web intelligence meets brain informatics, 36–58, Springer

Liao QV, Sundar SS, Lum K, Kearns M, Wu AX, Borgesius FZ (2022) Designing for responsible trust in AI systems: a communication perspective. In Lum K, Kearns M, Wu AX, Borgesius FZ (eds), Proceedings of the 2022 ACM conference on fairness, accountability, and transparency, 1257–1268 ACM

Lin W-H, Jin R, Hauptmann A, Zhong N, Liu J, Yao Y (2003) Web image retrieval re-ranking with relevance model. In Zhong N, Liu J, Yao Y (eds), Proceedings IEEE/WIC international conference on web intelligence (WI 2003), 242–248, IEEE

Lin Y-R, Sundaram H, Chi Y, Tatemura J, Tseng BL, Li Y, Raghavan VV (2007) Blog community discovery and evolution based on mutual awareness expansion. In Li Y, Raghavan VV (eds), IEEE/WIC/ACM international conference on web intelligence (WI'07)

Lin KH-Y, Yang C, Chen H-H, Zhong N, Liu J, Yao Y (2008) Emotion classification of online news articles from the reader's perspective. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 1, 220–226, IEEE

Liu J (2001) Autonomous agents and multi-agent systems. World Scientific, Singapore

Liu J, Jing H, Tang Y (2002) Multi-agent oriented constraint satisfaction. Artif Intell 136:101–144

Liu J, Wong CK, Hui KK (2003) An adaptive user interface based on personalized learning. IEEE Intell Syst 18:52–57

Liu J, Gottlob G, Walsh T (2003) Web Intelligence (WI): what makes wisdom Web? In Gottlob G, Walsh T (eds) Proceedings of the 18th international joint conference on artificial intelligence, IJCAI'03, 1596-1601. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA

Liu B, Ma Y, Yu PS, Zhong N, Liu J, Yao Y (2003) Discovering business intelligence information by comparing company web sites. In Zhong N, Liu J, Yao Y (eds) Web Intelligence, 105–127, Springer

Liu J, Zhang S, Ye Y, Zhong N, Liu J, Yao Y (2003) Agent-based characterization of web regularities. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 19–36, Springer

Liu J, Jin X, Tang Y (2004) Multi-agent collaborative service and distributed problem solving. Cogn Syst Res 5:191–206

Liu J, Birnbaum L, Li Y, Raghavan VV (2007) Measuring semantic similarity between named entities by searching the web directory. In Li Y, Raghavan VV (eds) IEEE/WIC/ACM international conference on web intelligence (WI'07), 461–465, IEEE

Liu J, Chen Y-W, Yang G-K, Lu Y-Z (2011) Self-organized combinatorial optimization. Expert Syst Appl 38:10532–10540

Liu X et al (2014) Data-driven composition for service-oriented situational web applications. IEEE Trans Serv Comput 8:2–16

Liu X, et al (2023) GPT understands, too. AI Open

Lnenicka M, Komarkova J (2019) Big and open linked data analytics ecosystem: theoretical background and essential elements. Gov Inf Q 36:129–144

Lu Y (2017) Industry 4.0: a survey on technologies, applications and open research issues. J Ind Inf Integr 6:1–10

Lu J, Wang Y, Liang J, Chen J, Liu J, Zhong N, Liu J, Yao Y (2008) An approach to deep web crawling by sampling. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 1, 718–724, IEEE

Lugrin B, Pelachaud C, Traum D (2022) The handbook on socially interactive agents: 20 years of research on embodied conversational agents, intelligent virtual agents, and social robotics volume 2: interactivity, platforms, application ACM

Macal CM (2016) Everything you need to know about agent-based modelling and simulation. J Simul 10:144–156

Maedche A, Pekar V, Staab S, Zhong N, Liu J, Yao Y (2002) Ontology learning part one-on discovering taxonomic relations from the web. In Zhong N, Liu J, Yao Y (eds) Web intelligence, 301–319, Springer

Mahmud R, Kotagiri R, Buyya R (2018) Fog computing: a taxonomy, survey and future directions. Springer, Cham

Mao Y, You C, Zhang J, Huang K, Letaief KB (2017) A survey on mobile edge computing: the communication perspective. IEEE Commun Surv Tutor 19:2322–2358

Martin P, Zhong N, Liu J, Yao Y (2003) Knowledge representation, sharing, and retrieval on the web. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 243–276 Springer

Massa P, Hayes C, Zhong N, Liu J, Yao Y (2005) Page-rerank: using trusted links to re-rank authority. In Zhong N, Liu J, Yao Y (eds), The 2005 IEEE/WIC/ACM international conference on web intelligence (WI'05), 614–617, IEEE

McRorie M et al (2012) Evaluation of four designed virtual agent personalities. IEEE Trans Affect Comput 3:311–322

Mehonic A, Kenyon AJ (2022) Brain-inspired computing needs a master plan. Nature 604:255–260

Mohammed AH, Ali AH (2021) Survey of BERT (bidirectional encoder representation transformer) types. J Phys: Conf Ser 1963:012173

Ngo CL, Nguyen HS, Zhong N, Liu J, Yao Y (2005) A method of web search result clustering based on rough sets. In Zhong N, Liu J, Yao Y (eds), The 2005 IEEE/WIC/ACM international conference on web intelligence (WI'05), 673–679 IEEE

Nie J-Y, Chen J, Zhong N, Liu J, Yao Y (2002) Exploiting the web as parallel corpora for cross-language information retrieval. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 218–239 Springer

Nishida T (2007) Conversational informatics and human-centered web intelligence. IEEE Intell Inform Bull 8:19–28

Niu L, Lu J, Chew E, Zhang G, Li Y, Raghavan VV (2007) An exploratory cognitive business intelligence system. In Li Y, Raghavan VV (eds) IEEE/WIC/ACM international conference on web intelligence (WI'07), 812–815 IEEE

Ortega-Bolaños R et al (2024) Applying the ethics of AI: a systematic review of tools for developing and assessing AI-based systems. Artif Intell Rev 57:110

Page MJ et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg 88:105906

Pinyol I, Sabater-Mir J (2013) Computational trust and reputation models for open multi-agent systems: a review. Artif Intell Rev 40:1–25
Postma M, Goedhart J (2019) PlotsOfData-a web app for visualizing data together with their summaries. PLoS Biol 17:e3000202
Pramanik MI, Lau RY, Demirkan H, Azad MAK (2017) Smart health: big data enabled health paradigm within smart cities. Expert Syst Appl 87:370–383
Qu Y, Pokhrel SR, Garg S, Gao L, Xiang Y (2020) A blockchained federated learning framework for cognitive computing in industry 4.0 networks. IEEE Trans Industr Inf 17:2964–2973
Rafferty J, Nugent CD, Liu J, Chen L (2017) From activity recognition to intention recognition for assisted living within smart homes. IEEE Trans Human-Mach Syst 47:368–379
Ren J, Guo H, Xu C, Zhang Y (2017) Serving at the edge: a scalable IoT architecture based on transparent computing. IEEE Network 31:96–105
Rigutini L, Maggini M, Liu B, Zhong N, Liu J, Yao Y (2005) An EM based training algorithm for cross-language text categorization. In Zhong N, Liu J, Yao Y (eds), The 2005 IEEE/WIC/ACM international conference on web intelligence (WI'05), 529–535, IEEE
Ritchie K (2020) Social structures and the ontology of social groups. Res 100:402–424
Saghezchi FB et al (2015) Drivers for 5G. Wiley, Hoboken
Sieg A, Mobasher B, Burke R, Li Y, Raghavan VV (2007) Ontological user profiles for representing context in web search. In Li Y, Raghavan VV (eds) 2007 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology-workshops, 91–94, IEEE
Sievert C (2020) Interactive web-based data visualization with R, plotly, and shiny. CRC Press, Boca Raton
Sisinni E, Saifullah A, Han S, Jennehag U, Gidlund M (2018) Industrial internet of things: challenges, opportunities, and directions. IEEE Trans Industr Inf 14:4724–4734
Speretta M, Gauch S, Zhong N, Liu J, Yao Y (2005) Personalized search based on user search histories. In Zhong N, Liu J, Yao Y (eds) The 2005 IEEE/WIC/ACM international conference on web intelligence (WI'05), 622–628, IEEE
Stronge AJ, Rogers WA, Fisk AD (2006) Web-based information search and retrieval: effects of strategy use and age on search success. Hum Factors 48:434–446
Takama Y, Saeed K (2007) Introduction of humatronics-towards integration of web intelligence and robotics. In Saeed K (ed) 6th International conference on computer information systems and industrial management applications (CISIM'07), 37–44, IEEE
Tan A-H IEEE (2002) Personalized information management for web intelligence. 2002 IEEE world congress on computational intelligence. 2002 IEEE international conference on fuzzy systems. FUZZ-IEEE'02. Proceedings (Cat. No. 02CH37291), Vol 2, 1045–1050, IEEE
Tan A-H, Ong H-L, Pan H, Ng J, Li Q-X (2004) Towards personalised web intelligence. Knowl Inf Syst 6:595–616
Tao X, Li Y, Zhong N (2010) A personalized ontology model for web information gathering. IEEE Trans Knowl Data Eng 23:496–511
Tao X et al (2020) Mining health knowledge graph for health risk prediction. World Wide Web 23:2341–2362
Tao X, Shaik TB, Higgins N, Gururajan R, Zhou X (2021) Remote patient monitoring using radio frequency identification (RFID) technology and machine learning for early detection of suicidal behaviour in mental health facilities. Sensors 21:776
Thelwall M (2005) Scientific web intelligence: finding relationships in university webs. Commun ACM 48:93–96
Tolosana R, Vera-Rodriguez R, Fierrez J, Morales A, Ortega-Garcia J (2020) Deepfakes and beyond: a survey of face manipulation and fake detection. Inform Fusion 64:131–148
Tomazela B, Hara CS, Ciferri RR, de Aguiar Ciferri CD (2013) Empowering integration processes with data provenance. Data Knowl Eng 86:102–123
Tsai C-W, Lai C-F, Chao H-C, Vasilakos AV (2015) Big data analytics: a survey. J Big Data 2:1–32
van Steen M, Tanenbaum AS (2016) A brief introduction to distributed systems. Computing 98:967–1009
Vlacheas P et al (2013) Enabling smart cities through a cognitive management framework for the internet of things. IEEE Commun Mag 51:102–111
Völker J, Rudolph S, Zhong N, Liu J, Yao Y (2008) Fostering web intelligence by semi-automatic OWL ontology refinement. In Zhong N, Liu J, Yao Y (eds) 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 1, 454–460, IEEE
Wang G et al (2024) Brain-inspired artificial intelligence research: a review. Sci China Technol Sci 67:2282–2296
Xi Z, et al (2023) The rise and potential of large language model based agents: a survey. arXiv preprint arXiv:2309.07864
Xia S, et al (2024) GBSVM: an efficient and robust support vector machine framework via granular-ball computing. IEEE Trans Neural Networks Learn Syst

Xu G, Zhang Y, Yi X, Zhong N, Liu J, Yao Y (2008) Modelling user behaviour for web recommendation using LDA model. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 529–532, IEEE

Yan Y, Qian Y, Sharif H, Tipper D (2012) A survey on smart grid communication infrastructures: motivations, requirements and challenges. IEEE Commun Surv Tutor 15:5–20

Yan Y, et al (2021) ConSERT: a contrastive framework for self-supervised sentence representation transfer. CoRR, abs/2105.11741

Yang J, Choi J, Zhong N, Liu J, Yao Y (2003) Knowledge-based wrapper induction for intelligent web information extraction. In Zhong N, Liu J, Yao Y (eds), Web intelligence, 153–172 Springer

Yang B, Cheung W, Liu J (2007) Community mining from signed social networks. IEEE Trans Knowl Data Eng 19:1333–1348

Yang B, Lei Y, Liu J, Li W (2016) Social collaborative filtering by trust. IEEE Trans Pattern Anal Mach Intell 39:1633–1647

Yang B, Pei H, Chen H, Liu J, Xia S (2016) Characterizing and discovering spatiotemporal social contact patterns for healthcare. IEEE Trans Pattern Anal Mach Intell 39:1532–1546

Yao Y (2018) Three-way decision and granular computing. Int J Approx Reason 103:107–123

Yao Y (2022) Symbols-meaning-value (SMV) space as a basis for a conceptual model of data science. Int J Approx Reason 144:113–128

Yao Y, SPIE (2006) Granular computing for data mining. In SPIE (ed), Data mining, intrusion detection, information assurance, and data networks security 2006, Vol 6241, 44–55, SPIE

Yao Y, Zhong N, Liu J, Ohsuga S, Zhong N, Yao Y, Liu J, Ohsuga S (2001) Web intelligence (WI): research challenges and trends in the new information age. In Zhong N, Yao Y, Liu J, Ohsuga S (eds), Web intelligence: research and development, pp 1–17 Springer, Berlin

Yong S, Hagenbuchner M, Tsoi AC, Zhong N, Liu J, Yao Y (2008) Ranking web pages using machine learning approaches. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 677–680, IEEE

Yu H, Shen Z, Leung C, Miao C, Lesser VR (2013) A survey of multi-agent trust management systems. IEEE Access 1:35–50

Yu C, Liu J, Nemati S, Yin G (2021) Reinforcement learning in healthcare: a survey. ACM Comput Surv (CSUR) 55:1–36

Yuan X, Lee J-H, Kim S-J, Kim Y-H (2013) Toward a user-oriented recommendation system for real estate websites. Inf Syst 38:231–243

Yue Y-G, He P (2018) A comprehensive survey on the reliability of mobile wireless sensor networks: taxonomy, challenges, and future directions. Inform Fusion 44:188–204

Zhai Z, Xu H, Jia P, Zhong N, Liu J, Yao Y (eds) (2008) Identifying opinion leaders in BBS. In Zhong N, Liu J, Yao Y (eds), 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, Vol 3, 398–401 IEEE

Zhang Q et al (2023) Information fusion for multi-scale data: survey and challenges. Inform Fusion 100:101954

Zhao J, Huang JX, Deng H, Chang Y, Xia L (2022) Are topics interesting or not? An LDA-based topic-graph probabilistic model for web search personalization. ACM Trans Inform Syst 40:1–24

Zheng Z, Xie S, Dai H-N, Chen X, Wang H (2018) Blockchain challenges and opportunities: a survey. Int J Web Grid Serv 14:352–375

Zhong N (2002) Representation and construction of ontologies for web intelligence. Int J Found Comput Sci 13:555–570

Zhong N, Yao Y, Ohsuga S, Zytkow JM, Rauch J (1999) Peculiarity oriented multi-database mining. In Zytkow JM, Rauch J (eds), Principles of data mining and knowledge discovery: third European conference, PKDD'99, Prague, Czech Republic, September 15-18, 1999. Proceedings 3, 136–146, Springer

Zhong N, Liu J, Yao Y, Ohsuga S, Matthew F, Miau C, SSY (2000) Web intelligence (WI). In Matthew FC, Miau SSY (eds) Proceedings 24th annual international computer software and applications conference. COMPSAC2000, pp 469–470. IEEE Computer Society

Zhong N et al (2013) Research challenges and perspectives on Wisdom Web of Things (W2T). J Supercomput 64:862–882

Zhong N et al (2015) Brain informatics-based big data and the wisdom web of things. IEEE Intell Syst 30:2–7

Zhong N, Liu J, Yao Y (2003) Web intelligence. Springer, Cham

Zhong N, et al (2004) Data mining for targeted marketing. Intelligent technologies for information analysis, pp 109–131

Zhong N, Liu J, Tan P-N, Kumar V (2004) Discovery of web robot sessions based on their navigational patterns. Intell Technol Inform Anal 193–222

Zhong N, Liu J, Yao Y (2007) Envisioning intelligent information technologies through the prism of web intelligence. Commun ACM 50:89–94

Zhong N, Liu J, Yao Y (2007) Web intelligence (WI). Wiley encyclopedia of computer science and engineering pp 1–11
Zhong N, Ma J, Liu J, Huang R, Tao X (2016) Wisdom web of things. Springer, Cham
Zhong N, Liu J, Shi Y, Yao Y (2018) An interview with professor Raj Reddy on web intelligence (WI) and computational social science (CSS). Web Intell 16:143–146
Zhou B, Yao Y (2010) Evaluating information retrieval system performance based on user preference. J Intell Inform Syst 34:227–248
Zorrilla M, Millan S, Menasalvas E, Hu X et al (2005) Data web house to support web intelligence in e-learning environments. In Hu X et al (eds), 2005 IEEE international conference on granular computing, Vol 2, 722–727, IEEE
Zou J, Huang J, Ren Z, Kanoulas E (2023) Learning to ask: conversational product search via representation learning. ACM Trans Inform Syst 41:1–27

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Hongzhi Kuai[1,2,8] · Jimmy X. Huang[3] · Xiaohui Tao[4] · Gabriella Pasi[5] · Yiyu Yao[6] · Jiming Liu[7] · Ning Zhong[8]

✉ Ning Zhong
zhong@maebashi-it.ac.jp

Hongzhi Kuai
kuaihz@cqupt.edu.cn

Jimmy X. Huang
jhuang@yorku.ca

Xiaohui Tao
Xiaohui.Tao@usq.edu.au

Gabriella Pasi
gabriella.pasi@unimib.it

Yiyu Yao
Yiyu.Yao@uregina.ca

Jiming Liu
jiming@hkbu.edu.hk

1 School of Computer Science and Technology/School of Artificial Intelligence, Chongqing University of Posts and Telecommunications, Chongqing, China

2 Key Laboratory of Big Data Intelligent Computing, Chongqing University of Posts and Telecommunications, Chongqing, China

3 School of Information Technology, York University, North York, ON, Canada

4 School of Mathematics, Physics, and Computing, University of Southern Queensland, Toowoomba, QLD, Australia

5 Department of Informatics, Systems, and Communication, University of Milano-Bicocca, Milano, Italy

6 Department of Computer Science, University of Regina, Regina, Canada

7 Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China

8 Faculty of Engineering, Maebashi Institute of Technology, Maebashi, Gunma, Japan