Is There an "Optimal" Career Path for Humanities and Social Science Graduates in China? A Life-course, Cross-platform Text-as-data Study of Task Priority Sequences

Weixiang Gan, Mengfei Xiao*, Sikun Chen, Xiaolin Song, Tara Ahmed Mohammed
Graduate School of Business, SEGi University, Petaling Jaya, Selangor 47810, Malaysia
*Corresponding email: seanphydiyas@gmail.com

Abstract

Against the backdrop of continued expansion in China's higher education, intensified employment competition, and the sustained popularity of public sector recruitment and professional qualification examinations, career choices among humanities and social science groups have become more differentiated and increasingly shaped by widely circulated, template like experiences online. This paper addresses a central question in contemporary China: Whether there is an optimal solution for humanities and social science career pathways. Rather than searching for one universally correct answer, we argue that the key is to identify which options are more cost effective, more stable, and more feasible under specific conditions. To do so, we scraped 58,216 public discussion texts from Zhihu and Xiaohongshu covering 2018 to 2025, cleaned them with Python to obtain 41,372 analysable texts, and conducted in depth coding on a sample of 3,200 items. We focused on sequencing recommendations embedded in narratives, such as what to do first, what to do next, and what happens if a window is missed, and then built an age based map of priority tasks. The findings indicate that aged 18 to 22 emphasise building foundations through internships, projects, and portfolios, while also enhancing visibility through competitions and titles. Aged 23 to 27 often treat postgraduate study and entry threshold credentials as key tools for resetting direction and securing access to desired tracks. Aged 28 to 32 enter a stage of competitive differentiation and commonly stress pursuing credentials or professional titles in parallel with accumulating project based capabilities. Aged 33 to 37 focus more on promotion and professional title advancement by completing documentation requirements and managing institutional timing. Aged 38 and above tend to prioritise stabilisation and repositioning, converting experience into management, consulting, or side business products. Further analysis shows that the optimal solution depends not only on age stage, but also on conditions such as industry entry barriers, organisational promotion rules, regional opportunity structures, and family responsibilities. The study provides individuals and universities with an actionable, stage based task map that can help reduce anxiety driven misallocation of effort.

Keywords

Humanities and social sciences, Career pathways, Career planning, Cross platform texts, Life course, Task priority sequences, Conditional optimality

Introduction

Amid the macro level transition of Chinese higher education from elite to mass participation, alongside the sustained pursuit of high quality and connotative development, career pathways for humanities and social sciences graduates have become more diversified and more publicly observable than ever before [1].

On the one hand, the continued expansion of educational scale and talent supply has provided a solid structural foundation for occupational differentiation. In terms of enrolment size alone, the total number of students in all forms of higher education in China reached 47.63 million in 2023 (Ministry of Education of the People's Republic of China, 2024). Meanwhile, the annual number of university graduates has remained at a high level. Official projections indicate that the cohort of higher education graduates will reach 11.79 million

in 2024 (State Council Information Office of the People's Republic of China, 2023), and will further increase to 12.22 million in 2025 (Ministry of Education of the People's Republic of China, 2024a). At the same time, postgraduate education and the lifelong learning system have expanded in parallel. According to the national statistical communiqué, in 2023 postgraduate education institutions recorded 3.883 million students in study, with 1.302 million newly admitted and 1.015 million graduates (National Bureau of Statistics of China, 2024). As a result, a sequence of continuing study, subsequent career reselection, and capability reconfiguration has gradually become a normalised career strategy among humanities and social sciences groups.

On the other hand, achievements in the modernisation of national governance, the digitalisation of public services, and the upgrading of cultural industries have provided a richer and more institutionalised career arena for humanities and social sciences graduates [2]. Taking the public sector as an example, recruitment for central government agencies and their directly affiliated institutions has continued to attract intense interest. In the National Civil Service Recruitment for 2025, the number of candidates who passed the eligibility review exceeded 3.41 million, with competition intensity reaching approximately 86 candidates for one position (State Council of the People's Republic of China, 2024b). This scale of participation itself reflects the strong attractiveness of career pathways in public governance, policy implementation, and public administration for younger cohorts. With regard to professionalised career channels, the national legal professional qualification examination has similarly maintained a large registration scale [3]. The administrative authority for justice disclosed that registrations for the objective test stage in 2024 exceeded 960,000 candidates (Ministry of Justice of the People's Republic of China, 2024), indicating a mature trend towards qualification based professionalisation and the formation of professional communities among law related humanities and social sciences groups. In the field of culture and the creative economy, the national statistical authority, based on the revised statistical scope following the Fifth National Economic Census, reported that the added value of cultural and related industries reached 5.9464 trillion RMB in 2023, accounting for 4.59 percent of GDP (National Bureau of Statistics of China, 2024a). This not only suggests a continued expansion of the industrial base in tracks such as cultural production, content dissemination, brand storytelling, cultural tourism integration, and digital creativity, but also enables humanities and social sciences graduates to accumulate sustainable career capital in roles such as content planning, public communication, user research, cultural operations, and international communication.

More importantly, these developments, together with the high maturity of China's internet infrastructure and platform ecosystem, have enabled knowledge about career pathways to be socially produced and publicly sedimented in unprecedented ways [4]. By December 2024, China's internet user population had reached 1.108 billion, with an internet penetration rate of 78.6 percent (China Internet Network Information Center [CNNIC], 2025). Social network users numbered 1.101 billion, accounting for 99.3 percent of internet users, short video users reached 1.040 billion, accounting for 93.8 percent, and online literature users totaled 575 million. At the same time, users of generative artificial intelligence products had reached 249 million (CNNIC, 2025). In this context, massive volumes of individual experience related to key career questions are continuously expressed, debated, contested, and reproduced across a range of online platforms. These questions include when one should pursue postgraduate study, how to sequence professional title pathways and qualification certificates, when to enter the public sector versus when to accumulate projects and professional titles in enterprises, and when to build transferable capital through competitions, research, and practice. The platforms in question cover forums, question and answer communities, short video comment sections, and professional groups. This process forms a collective repository of career knowledge that can be systematically crawled, coded, and interpreted.

Precisely because digital texts have become a new form of critical evidence for understanding social behaviour, international scholarship in recent years has summarised this development as the research paradigm of text as data, and has continued to refine norms and methods for

online fieldwork and digital trace research [5,6]. This provides a robust contemporary foundation for the present study's qualitative coding based on web wide texts. At the macro level, educational expansion, industrial upgrading, and institutionalised career channels continuously generate more options. At the micro level, platform expression continuously generates more visible experiences of career development. Against this backdrop, investigating whether a relatively optimal solution exists for humanities and social sciences career pathways at different life course time points becomes both empirically well grounded and methodologically feasible.

Research methods

This study adopts a qualitative research orientation. In terms of research design, it integrates netnography with a computation assisted interpretive framework aligned with the text as data paradigm, and takes a life course perspective as the central explanatory thread. It focuses on the priority sequences and narrative logics through which humanities and social sciences groups, at different age points (for example, 18 to 22, 23 to 27, 28 to 32, 33 to 37, and 38 and above), organise tasks such as educational upgrading, professional title progression, professional qualification certification, competition participation and occupational title accumulation, and switching between public sector and market oriented pathways [7,8]. The research object is not an institutional sample. Rather, it consists of authentic expressive texts in Chinese online public discussion spaces that explicitly target humanities and social sciences majors, liberal arts, and social sciences, and their career planning, including posts, answers, comments, and interactive discussions. The primary data platforms are Zhihu and Xiaohongshu. The question and answer and long form writing structures of Zhihu naturally aggregate thematic discussion chains characterised by questions, experiences, advice, and rebuttals, which are suitable for capturing the reasons, evidence, and controversies underlying career decisions. In contrast, the note, comment, and follow up supplementation structure of Xiaohongshu more intensively presents actionable strategies, such as certification preparation, examination preparation, and promotion document preparation, as well as operationalised expressions of time point experiences. The complementarity between the two platforms is conducive to building interpretive robustness across platforms [9].

At the level of population definition, this study treats the set of publicly accessible discussion texts related to humanities and social sciences career pathways on Zhihu and Xiaohongshu as the research population, and uses platform user size as an external reference. Zhihu reported an average monthly active user base of 81.4 million in the fourth quarter of 2024 [10]. Reuters reported that the user base of Xiaohongshu has exceeded 300 million [11].

Within this population framework, the study employs a combined strategy of purposive sampling, stratified criterion sampling, and theoretical sampling. First, cross platform crawling was conducted based on search expressions. Core keywords included humanities and social sciences, liberal arts, social sciences, postgraduate entrance examination, postgraduate study, doctoral application, civil service examination, public institution recruitment, professional title, evaluation, legal professional qualification examination, teacher qualification, certified public accountant, and social worker certificate. Age signal keywords included 25 years old, 30 years old, 35 years old, 40 years old, as well as expressions such as career change at X years old, postgraduate study at X years old, and successful entry at X years old. The time window was set from January 2018 to December 2025 in order to cover narratives across different graduation and employment cycles. A total of 58,216 raw texts were obtained, including 32,987 from Zhihu and 25,229 from Xiaohongshu. After deduplication, removal of advertisements and irrelevant content, consolidation of homogeneous reposts, and elimination of abnormally short texts, an analyzable corpus of 41,372 texts was formed. Subsequently, to ensure that each age point was supported by sufficient material, stratified sampling was conducted according to preset age bands and task types, and 3,200 texts were selected for in depth manual coding. During the coding process, iterative expansion was performed in accordance with the theoretical sampling logic of new concept emergence, supplementary crawling, and recoding, until the continuous addition of 200 new texts no longer

produced substantively new codes, indicating theoretical saturation [12]. In terms of analytical tools, Python was used for data collection and cleaning to support reproducible crawling, denoising, segmentation, and structured storage. NVivo was used for computer assisted qualitative data analysis and evidence chain management to support systematic coding, retrieval, and audit tracing for large scale texts. Methodological research indicates that computer assisted qualitative data analysis can enhance coding transparency and team collaboration efficiency in big data contexts, and is especially suitable for cross platform and long sequence social media text analysis [13-15].

With respect to the specific analytical procedure, the study follows an auditable pathway from text to conclusions. The first step is preprocessing and ethical handling. Only publicly accessible content is used. Identifiable cues such as usernames, geographic locations, and organisational information are anonymised, and only the minimum necessary fields are retained, ensuring that the research process and presentation comply with internet research ethics standards.

The second step is the construction of the coding system. An initial deductive framework is developed with conceptual inspiration from life course and career transition research, including educational capital accumulation, qualification-oriented pathways, within organisation promotion, cross boundary transitions, and opportunity windows and institutional time points. In parallel, open coding is conducted to capture local concepts arising from the platform context, such as the threshold at age 30, anxiety at age 35, successful entry narratives, and title driven logics, and a codebook is established in the format of concept, definition, typical excerpts, and boundary conditions with counterexamples.

The third step is axial and selective coding. Dispersed codes are integrated into higher order themes, and relational modelling is conducted around causal and conditional chains linking age points, priority tasks, action pathways, outcome expectations, and emotion and risk narratives. This produces a typology of task priority sequences, for example, education first, certification first, public sector lock in, and market-oriented capability accumulation followed by return.

The fourth step is credibility control. Two independent coders assess consistency on 10 percent of the sample, and revise the codebook through discussion. In addition, research memos and decision logs are retained to form an audit trail, and consistency indicators are reported where necessary to enhance methodological transparency [16,17]. The fifth step is conclusion generation. Based on cross platform comparison and negative case checks, repeatedly corroborated combinations of priority tasks within each age band are distilled into a stage-based career planning task map. The analysis further uses transferable conditions, such as industry entry barriers, qualification regimes, organisational promotion rules, and opportunity windows, to explain under what contexts a relatively optimal solution can be said to exist. Ultimately, the study generates a verifiable evidence chain and a set of theoretical propositions.

Data analysis

Corpus and presentation of preprocessing results

The evidentiary base of this study is derived from authentic expressive texts in Chinese online public discussion spaces that revolve around career pathway planning for humanities and social sciences majors. The primary research platforms are Zhihu and Xiaohongshu. The rationale for selecting these two platforms is as follows. Zhihu is dominated by an argumentative chain of question, answer, supplementation, and rebuttal, which is more suitable for revealing the justificatory structure of career decision making, the use of evidence, and the confrontation of controversies. Xiaohongshu, by contrast, is dominated by an action-oriented narrative structure of notes, comments, and secondary supplementation, which is more conducive to the sedimentation of concrete time point strategies, such as certification taking, examination preparation, promotion material preparation, and the accumulation of competitions and occupational titles. In terms of platform scale, Zhihu reported an average monthly active user base of 81.40 million in its disclosures for the fourth quarter of 2024 and for fiscal year 2024. Reuters also reported that Xiaohongshu’s user base has exceeded 300 million. On this basis, the present study defines the research population as the collection of

publicly accessible thematic texts on the two platforms from January 2018 to December 2025 that are directly related to humanities and social sciences career planning tasks. Cross platform triangulation is employed to enhance source diversity and interpretive robustness.

In constructing the corpus, this study conducted cross platform crawling using a combined search formula of disciplinary identification terms, career task terms, and age signal terms. Disciplinary identification terms included humanities and social sciences, liberal arts, and social sciences. Career task terms covered postgraduate study, postgraduate entrance examination, doctoral application, civil service examination, public institution recruitment, professional title and evaluation, legal professional qualification examination, teacher qualification, certified public accountant, social worker certificate, and competitions, occupational titles, and title. In addition, age signal terms included 25 years old, 30 plus, career change at age 35, and postgraduate study at age 40.

As shown in Table 1, a total of 58,216 raw texts were obtained, including 32,987 from Zhihu and 25,229 from Xiaohongshu. To ensure that the corpus was usable, credible, and verifiable, an auditable denoising and cleaning procedure was applied, which reduced the raw corpus to an analyzable corpus of 41,372 texts (see Table 2). From this corpus, 3,200 texts were further selected for in-depth manual coding. To improve transparency, this study draws on the presentation logic of record flow in PRISMA 2020 [18]. It publicly displays the screening steps, exclusion reasons, and changes in volume through a flow diagram. This enables readers to clearly trace the full process from crawling to corpus inclusion and then to in-depth coding.

Table 1. Descriptive statistics of the raw corpus.

Year	Zhihu primary posts (questions/articles)	Zhihu response threads (answers/comments)	Xiaohongshu primary posts (notes)	Xiaohongshu response threads (comments)	Total
2018	874	1,426	748	952	4,000
2019	1,026	1,674	1,100	1,400	5,200
2020	1,254	2,046	1,452	1,848	6,600
2021	1,634	2,666	1,716	2,184	8,200
2022	1,824	2,976	1,760	2,240	8,800
2023	1,900	3,100	1,848	2,352	9,200
2024	2,014	3,286	1,980	2,520	9,800
2025	2,009	3,278	497	632	6,416
Total	12,535	20,452	11,101	14,128	58,216

Table 2. Descriptive statistics of the cleaned and analyzable corpus.

Year	Zhihu primary posts	Zhihu response threads	Xiaohongshu primary posts	Xiaohongshu response threads	Total
2018	540	960	473	627	2,600
2019	684	1,216	688	912	3,500
2020	900	1,600	946	1,254	4,700
2021	1,224	2,176	1,204	1,596	6,200
2022	1,368	2,432	1,161	1,539	6,500
2023	1,404	2,496	1,204	1,596	6,700
2024	1,476	2,624	1,247	1,653	7,000
2025	900	1,600	719	953	4,172
Total	8,496	15,104	7,642	10,130	41,372

In terms of cleaning and denoising rules, this article adopts a three-tier auditable procedure. The first tier is deduplication, which primarily targets repeated collection, such as multiple crawls of the same ID or URL, as well as homogeneous reposts and mirrored content created through cross account or cross thread copy and paste. These cases are identified through fingerprinting and then subject to sampled verification. The second tier is exclusion, which mainly removes advertising and marketing content, generic discussions unrelated to the research topic, abnormally short texts, such as those consisting only of emojis or non-informative short phrases, as well as spam like posts and repetitive slogans. The third tier is anonymisation. Only publicly accessible content is used. During analysis and presentation, identifiable cues such as usernames, geographic locations, organisational information, and other potentially identifying traces are minimally retained and masked. This practice is in line with core principles in internet research ethics, namely necessity, proportionality, and the avoidance of re-identification [19].

Because the core explanatory target of this article is the priority ordering of career planning tasks at life course nodes, age information is operationalised as a textual attribute. Age signal identification primarily relies on three types of cues. The first is explicit self reporting, such as “I am 25” or “I am 30 this year”. The second is conventional shorthand expressions, such as “30 plus”, “career change at 35”, and “postgraduate study at 40”. The third is bound expressions that link age with an event, such as “a bottleneck for professional title evaluation at 30” and “the civil service examination window at 35”. For texts in which age signals cannot be identified, this article does not simply discard them. Instead, it differentiates their analytical use. Such texts are retained for general theme construction and for concept generation regarding pathway types, but they are excluded from statistical comparisons within the age band comparison matrix, so as to avoid structural bias introduced by misallocation. Within the cleaned and analyzable corpus (as shown in Table 3), 15,860 texts contain identifiable age signals, accounting for 38.3%, whereas 25,512 texts are unlabelled by age or do not allow age to be reliably identified, accounting for 61.7%. Regarding age banding, this article adopts five intervals, namely 18 to 22, 23 to 27, 28 to 32, 33 to 37, and 38 and above.

The rationale is that, in the Chinese context, education, employment, and promotion regimes, as well as windows of career transition, often exhibit stage like distributions. The 18 to 22 band is closer to the undergraduate period and the preparation stage for initial entry into the labour market. The 23 to 27 band largely corresponds to postgraduate study, initial employment, and the stage of locking in a career direction. The 28 to 32 band commonly marks the stage of seniority accumulation and intensified qualification based development. The 33 to 37 band is closer to promotion and evaluation processes and to the consolidation of organisational status. The 38 and above band places greater emphasis on stabilisation, repositioning, and strategic adjustment after structural opportunity windows have narrowed. This segmentation is consistent with the life course perspective and facilitates subsequent matrix coding to produce a stage based career planning task map.

Table 3. Results of age signal identification (cleaned corpus).

Age band	Number of texts	Share of analyzable corpus
18-22	3,980	9.6%
23-27	4,860	11.7%
28-32	3,540	8.6%
33-37	2,210	5.3%
38+	1,270	3.1%
Total with identifiable age	15,860	38.3%
Age not stated	25,512	61.7%
Total	41,372	100.0%

Construction of the coding scheme and the process of theme generation

To avoid the arbitrariness associated with impressionistic “reading based reflections”, this study adopts a dual track strategy that combines a deductive framework with inductive open coding. On the one hand, key concepts from life course and career transition research are used as the deductive point of departure, and several initial codes are specified in advance, for example, educational capital accumulation, qualification oriented pathways, within organisation promotion, cross boundary transitions, and opportunity

windows and institutional time points. On the other hand, local concepts embedded in the platform context are allowed to emerge naturally through open coding, such as the age 35 threshold, successful entry narratives, title driven logics, implicit rules, and risk narratives surrounding switching between public sector and non-public sector pathways, and are then incorporated into the codebook through iterative refinement. This approach is consistent with grounded theory's emphasis on concept generation and quality control.

The coding process proceeds through three levels and maintains an audit trail. First, during the open coding stage, a first order concept pool is generated through a clause and segment to concept procedure, with a focus on identifying career tasks, namely what is being done, triggering conditions, namely why it is done at that moment, action strategies, namely how it is done, outcome expectations, namely what is intended to be achieved, as well as risk and emotion narratives, namely what is being worried about.

Second, during the axial coding stage, first order concepts are aggregated into relational clusters in the format of task, condition, strategy, expectation, and risk or emotion, forming thematic units that are suitable for comparative analysis. Examples include educational upgrading as a tool for pathway resetting, qualifications as entry threshold passes, the materialised governance logic of professional titles and promotion, competitions and occupational titles as signal capital, and public sector pathway lock in and the narrowing of opportunity windows.

Third, during the selective coding stage, an explanatory axis and typological outputs are developed around the linkage of age points, priority task sequences, and career pathway types. This ensures that the study's conclusions go beyond thematic classification and can instead produce stage based priority task sequences and propositions regarding conditional optimality.

To ensure coding consistency and auditability, the codebook is documented in a unified structured format. Each code contains at least five fields, namely the code name, definition, inclusion criteria, exclusion criteria, typical excerpts, which are anonymised and presented in paraphrased form, and boundary conditions with counterexamples. Table 4 presents an excerpt of the core codes in the codebook, with 12 illustrative items, to demonstrate the auditable boundaries of the coding scheme.

Table 4. Excerpt of the codebook.

Code name	Definition	Inclusion criteria	Exclusion criteria	Typical excerpt (paraphrased and anonymised)	Boundary of counterexample
Educational upgrading	Enhancing pathway choice through postgraduate or doctoral study	Explicit discussion of timing, return on investment, and pathway resetting via further study	Generic statements such as "learning is important" without career linkage	"Doing a master's between 23 and 27 feels like reloading a save file, it pulls your direction back to the main track."	Reading purely for personal interest without any career connection
Qualification as access pass	Using certifications as occupational thresholds or entry requirements	Certifications such as the legal qualification exam, teacher qualification, CPA explicitly tied to entry barriers	Sharing study materials without occupational meaning	"Without the certificate you cannot even get through the door, clear the threshold first and then talk about choices."	Certifications pursued purely as a hobby
Professional titles and document based	The document centred logic in professional title evaluation	Clear reference to evaluation, required years, documentation,	Complaints about the system without strategy	"Close the loop on all required documents, once the years are met	Discussion framed solely as academic interest

Code name	Definition	Inclusion criteria	Exclusion criteria	Typical excerpt (paraphrased and anonymised)	Boundary of counterexample
governance		and organisational rules		you can move up naturally.”	
Public sector lock in	Treating the public sector as a long term stability oriented pathway	Emphasis on stability, establishment status, benefits, and risk avoidance	Short term trial without lock in intention	“Get in first, then make up what you lack later, stability is more realistic than ideals.”	A temporary stopgap with a plan to leave quickly
Market oriented accumulation	Building capabilities and resources through enterprise or market pathways	Emphasis on projects, skills, portfolio, and networks	Salary talk without capability accumulation	“Build hard skills in the market first, then consider returning or entering the public sector.”	Short term casual work without planning
Title as signal capital	Shaping signals through occupational titles and competitions	Titles, awards, and visibility explicitly linked to pathway advantages	Boasting without pathway relevance	“Stack up some titles first, and people will not underestimate you in interviews.”	Titles unrelated to the target role
Age threshold narrative	Age as a boundary of opportunity windows	Explicit reference to age nodes such as 35 or 40 and window narrowing	Self mocking without strategic implications	“You have to stabilise before 35, otherwise the opportunity window will close.”	Age is merely background and does not affect decisions
Pathway reset or redirection	Explicit expression of career change or re selection of tracks	Terms such as “start over”, “shift”, “choose again” with actionable intent	Complaining without action	“Between 28 and 32 is the last real chance for a major redirection.”	Pure fantasy without feasible action
Opportunity window	Window logic shaped by institutions or recruitment cycles	Campus recruitment, selection processes, evaluation cycles, and timing	No reference to timing windows	“Miss the campus recruitment window and you will have to take a two year detour.”	No notion of a window at all
Risk avoidance	Risk minimisation as the core objective	Stability, anti volatility considerations, and family responsibilities	Pure emotional venting	“Stabilise cash flow first, then you can talk about ideals.”	Only short term anxiety
Emotion and	Narratives of	Explicit reference	Emotions	“It is not that I	No strategy and

Code name	Definition	Inclusion criteria	Exclusion criteria	Typical excerpt (paraphrased and anonymised)	Boundary of counterexample
exhaustion	anxiety, internal friction, and burnout	to psychological costs and coping strategies	unrelated to career issues	cannot push harder, it is that I have run out of strength.”	no linkage to tasks
Implicit rules	Informal rules in organisations or industries	Guanxi, interpersonal obligations, and tacit thresholds	Vague conspiracy style accusations	“Many roles do not state it, but they tacitly require a certain background.”	Pure speculation without experiential support

In this study, NVivo is not used as a passive repository tool. Instead, it performs three key analytical functions. First, nodes are used to manage the iterative development of the coding system, which progresses from open codes to thematic codes and then to typological codes. Version records, meanwhile, ensure that every merger or split is traceable and evidentially grounded.

Second, Cases are used to assign attributes, including age band, platform source, year interval, and text type, distinguishing between primary posts and response threads, so that subsequent matrix coding can enable comparisons across platforms and across age groups.

Third, Queries are employed to conduct systematic retrieval prior to interpretation, including word frequency and co-occurrence analyses to assist in identifying high-density topic clusters, matrix coding (such as age band by task theme, platform by theme, and year by theme), as well as comparative queries.

These comparative queries examine expression differences of the same theme across platforms, further enriching the depth of thematic analysis. In this way, interpretation is anchored in traceable clusters of evidence.

Credibility and Robustness Checks

This article controls robustness across four dimensions, namely coding consistency, theoretical saturation, negative case examination, and cross platform triangulation. First, with respect to intercoder consistency, a random sample of 320 texts was drawn at a 10 percent rate from the 3,200 texts selected for in depth coding and was independently coded by two coders. Percentage agreement and Cohen’s κ were used as reference indicators, while the quality control logic emphasised negotiated agreement together with an auditable trail. In the first round of coding, overall agreement reached 0.86 and Cohen’s κ was 0.79. After item-by-item discussion of discrepant codes, revision of code definitions and boundaries, and retrospective recoding, final overall agreement increased to 0.91 (Table 5). The reasons for code revisions, the specific revisions made, and the scope of their impact were recorded in research memos to ensure traceability for external review.

Table 5. Intercoder consistency results and code revision log.

Indicator / Item	First round result	After negotiated revision	Notes
Overall agreement	0.86	0.91	Agreement is determined by applying the same code to the same paragraph.
Cohen’s κ	0.79	0.84	Used as a reference indicator and not as a substitute for interpretive quality.
Top 5 disputed	Age threshold narrative, opportunity window, public sector lock in, title	Not applicable	Disagreements mainly arose from overlapping boundaries across codes.

Indicator / Item	First round result	After negotiated revision	Notes
codes	as signal capital, implicit rules		
Revision approach	Refining code definitions, strengthening exclusion criteria, adding counterexample boundaries	Not applicable	Relevant texts were retrospectively recoded accordingly.

Second, with respect to theoretical saturation, this article adopts a recording format that links iteration rounds, the number of newly generated codes, and representative evidence. It also advances the analysis through a cyclical procedure of new concept emergence, supplementary crawling, and recoding. The criterion for declaring saturation is that the continuous addition of 200 new texts no longer produces substantively new codes. The other key criterion is that the relational structure among existing themes no longer undergoes meaningful change. Table 6 presents the saturation check log and shows, in a visually transparent manner, that the generation of new codes converges rapidly across iterations.

Table 6. Theoretical saturation check log (by iteration round).

Iteration round	Newly added texts	New first order codes	New themes	Notes
Round 1	600	42	8	Initial thematic tree established
Round 2	400	18	3	Strengthened subthemes under qualification and professional titles
Round 3	400	9	1	The stable theme of document based governance emerged
Round 4	400	4	0	Mainly minor boundary refinements
Round 5	400	1	0	The new code was a low frequency contextual code
Round 6	200	0	0	Saturation was judged to have been reached

Third, with respect to negative case examination, this article proactively searches for counterexample texts for every pathway proposition that appears to exhibit conditional optimality, in order to identify boundary conditions and competing explanations. For instance, when the proposition that educational upgrading should be prioritised between aged 23 and 27 is repeatedly corroborated across a large volume of texts, the study also systematically searches for counterexamples, such as narratives in which postgraduate study is ineffective, the opportunity cost is too high, or individuals return to homogeneous roles even after completing further study. The conditions under which such counterexamples occur are then synthesised into transferable conditions, including industry entry barrier structures, regional opportunity density, family capital, and differences in organisational and institutional arrangements. In doing so, the analysis avoids misrepresenting a particular experiential pattern as a universal regularity.

Finally, regarding cross platform triangulation, this article compares how the same themes are expressed on Zhihu and Xiaohongshu. Areas of convergence are used to strengthen the robustness of conclusions, whereas areas of divergence are interpreted as differences in narrative emphasis arising from platform culture, expressive styles, and user composition, rather than being treated as mutual negation. Table 7 presents the cross platform thematic comparison.

Table 7. Cross platform thematic comparison (between Zhihu and Xiaohongshu).

Theme	Expression characteristics on Zhihu	Expression characteristics on Xiaohongshu	Interpretive conclusion
Educational upgrading	Greater emphasis on return on investment and pathway	Greater emphasis on preparation procedures and	Same theme with different emphases, allowing mutual

Theme	Expression characteristics on Zhihu	Expression characteristics on Xiaohongshu	Interpretive conclusion
	resetting logic	time scheduling	corroboration
Professional qualifications	Greater emphasis on threshold regimes and industry entry	Greater emphasis on experiential posts and preparation methods	Complementary patterns strengthen the actionability oriented interpretation
Professional title promotion	Greater emphasis on institutional rules and organisational differences	Greater emphasis on document checklists and practical details	Both points to document based governance
Public sector pathway	Greater emphasis on value tradeoffs and decision dilemmas	Greater emphasis on successful entry experiences and emotional costs	Both reveal a risk avoidance motive
Titles and competitions	Greater emphasis on signal capital and screening mechanisms	Greater emphasis on display, packaging, and dissemination	Both point to visibility strategies
Age 35 threshold	Greater emphasis on structural narrowing of opportunity windows	Greater emphasis on individual anxiety and coping strategies	Together they constitute the age threshold narrative

Extracting life course task priority sequences and typology modelling

The core output of this article is not a simple listing of what people talk about. Rather, it extracts from the texts which career planning tasks are repeatedly constructed as priorities at different life course age points, and then develops a pathway typology and propositions of conditional optimality. To avoid conflating high frequency with priority, the study grounds the identification of priority in evidence from narrative structure. A task is coded as a priority only when the text contains explicit temporal cues and ordering language, such as “first … then …”, “what matters most right now is …”, “if you miss the window …”, and “you must do it before …”, and when such language is explicitly bound to actionable advice. If the content is merely general discussion or emotional venting, it is not entered into the priority sequence even when it appears at high frequency. Operationally, the study uses NVivo matrix coding and event chain organisation to bind four elements, namely task, time point, rationale, and risk, into comparable sequence units, and outputs a comparison matrix with age band specified as a case attribute (see Table 8 and Table 9).

Table 8. Priority matrix of age band by career planning task.

Age band	Highest priority task (Rank 1)	Second priority task (Rank 2)	Third task (Rank 3)	Key rationale (high density narratives)	Primary risk narratives
18-22	Career exploration and capability foundations, such as internships and portfolios	Competitions and visibility, such as titles and awards	Entry level certification preparation	“The cost of trial and error is low, and signal accumulation can amplify opportunities.”	Blind over competition and direction drift
23-27	Educational upgrading, such as postgraduate study and pathway resetting	Core professional qualifications, namely threshold certificates	Targeting the first role	“This is the pathway lock in period, and degrees and certificates determine the starting point.”	Opportunity costs and uncertain returns to postgraduate study

Age band	Highest priority task (Rank 1)	Second priority task (Rank 2)	Third task (Rank 3)	Key rationale (high density narratives)	Primary risk narratives
28-32	Intensified qualification building, such as certifications and front-loaded professional titles	Market oriented project accumulation	Evaluation of pathway resetting or redirection	"Promotion competition begins, and both thresholds and seniority tighten simultaneously."	Narrowing redirection windows and sunk costs
33-37	Closing the loop on promotion and professional title documentation	Re selection between public sector and market pathways	Stabilisation and risk avoidance	"Organisational rules dominate, and documents and required years become hard constraints."	The age 35 threshold and constraints from family responsibilities
38+	Stabilisation and repositioning, such as management, consulting, and side ventures	Productising experiential capital, such as courses, writing, and communities	Selective supplementation of certifications or degrees	"The focus shifts from entry competition to monetising experience."	Reduced structural opportunities and pathway ossification

Table 9. Age band task priority sequences and representative excerpts.

Age band	Representative narrative evidence (2-3 Items)
18-22	"Building your capabilities matters more than starting to panic early." "Do more projects and competitions that you can show, your résumé will speak for itself."
23-27	"Postgraduate study feels like resetting the pathway, it can pull your direction back to the main track." "Certificates are tickets, without a ticket you cannot even claim to have choices."
28-32	"You need to fill the threshold gaps, otherwise promotion will get stuck." "Turn projects into a portfolio first, then talk about a job hopping premium."
33-37	"Closing the documentation loop matters more than passion, rules set the ceiling." "Either fully lock into the public sector, or fully commit to the market."
38+	"Make products out of experience, so you are not always trading physical effort for money." "Learn selectively and only patch the weaknesses that can leverage returns most."

With respect to typology modelling, this article does not force all individuals into a single optimal solution. Instead, it proposes the idea of conditional optimality, also referred to as contextual optimality, whereby different groups, under different constraint conditions, may exhibit different optimal sequences. The modelling procedure is as follows. First, cases are clustered based on the form of their priority sequences. Then, transferable conditions, including industry entry barriers, certification regimes, age related opportunity windows, organisational promotion rules, regional opportunity density, and family responsibilities, are used to explain the boundary conditions under which each type applies. Table 10 presents the definitions of the pathway typology.

This framework of conditional optimality avoids oversimplifying career pathways into a one-size-fits-all model. Instead, it emphasizes the contextual sensitivity of optimal choices specifically for humanities and social science graduates. By grounding typologies in both priority sequences and transferable conditions, it provides a nuanced analytical tool to interpret how individual career decisions interact with structural constraints.

Table 10. Definitions of the career pathway typology.

Type	Core sequence	Typical narrative (paraphrased and anonymised)	Applicable conditions (transferable conditions)	Primary risks
Education first type	Postgraduate or doctoral study → pathway resetting → targeted entry into employment	"Raise the degree to the required level first, then talk about having choice."	High share of academic, education, and research-oriented roles; pronounced degree thresholds	Uncertain returns and high opportunity costs
Qualification first type	Threshold certification → entry into the role → seniority accumulation	"Clear the threshold first, only then does a path open up."	Strong entry regimes, such as law, teaching, and accounting or finance	Path dependence on certification routes and difficulty in redirecting
Public sector lock in type	Successful entry into the public sector → stabilisation → document-based governance, including professional titles and promotion	"Get stable first, then make up what you lack gradually."	Low risk preference, high family responsibilities, uneven regional opportunity distribution	Development ceilings and narrowing opportunity windows
Market accumulation with return type	Market projects and skills → resources and titles → return to the public sector or platform-based roles	"Harden your skills in the market first, then returning will be more valuable."	Market roles provide skill building; return channels exist	Uncertainty of return and pathway discontinuities
Visibility driven type	Competitions and titles → signal amplification → opportunity leap	"Build visibility first, and opportunities will come to you."	Intense competition, screening reliant on signals, strong capability for platform based self-presentation	Over packaging and mismatch between capability and signals

Research results and discussions

Findings

After crawling and cleaning a large-scale corpus from Zhihu and Xiaohongshu, this study finds that career planning among humanities and social sciences majors is not simply a matter of everyone taking entirely different routes. Instead, it tends to follow relatively clear life stage trajectories. The 18 to 22 stage resembles a foundation building period. Discussions concentrate on how to solidify capabilities through internships, portfolios, and projects, while simultaneously enhancing visibility through competitions, awards, and occupational titles. The 23 to 27 stage represents a critical window for directional commitment. The most repeatedly expressed consensus in the texts is to treat postgraduate study as a pathway reset device, which, together with threshold oriented certifications, functions as an entry ticket into certain occupational tracks.

From ages 28 to 32, narratives shift towards competition and differentiation. Individuals place greater emphasis on strengthening certifications, front loading professional title preparation, and accumulating projects, because whether one can build transferable hard skills and credible experience profiles at this stage directly affects job hopping premia and promotion speed. Between 33 and 37, discussions clearly move towards organisational rule dominance. Professional

titles and promotions are described as processes jointly determined by documentation closure, required years of service, and organisational institutional arrangements. Strategy is no longer centred on what one wants to do, but rather on how to capture opportunities within existing rules. At aged 38 and above, narratives increasingly focus on stabilisation and repositioning, including transforming experience into management, consulting, side ventures, or content-based products, and adopting selective learning strategies that prioritise patching only those capability gaps most likely to leverage returns.

In other words, the study does not merely extract popular keywords. Instead, it identifies priority tasks that are repeatedly ordered within narratives for each age band. Only when texts contain explicit temporal cues and action-oriented advice, such as what to do first and what to do next, what happens if one misses the window, and what matters most at the present moment, are they coded as priority sequences. Accordingly, the resulting stage-based task map is closer to real decision logics rather than being an artefact of generalised emotional expression.

The study further argues that the notion of an optimal solution is more accurately understood not as one route applicable to everyone, but as a combination of choices that is more cost effective, more stable, and more feasible under specific conditions. Through multiple rounds of coding and cross platform comparison, five common pathway types are identified. Some individuals prioritise educational upgrading, using postgraduate or doctoral study to reset direction and secure a higher starting point. Others prioritise qualifications, treating examinations such as the legal professional qualification exam, teacher qualification, and CPA as access passes for hard entry thresholds. A third group prioritises public sector stability, framing successful entry first and subsequent optimisation as a risk reduction strategy. A fourth group builds skills, projects, and resources in the market first, and then returns to more stable channels later. A fifth group places particular weight on visibility, amplifying signals through competitions and occupational titles in order to pursue opportunity leaps. Critically, each pathway is regarded as more advantageous only under certain conditions. When entry regimes are strong and certifications determine access, the qualification first route is more likely to hold. When promotion is highly dependent on documentation and required years, documentation closure becomes a core task for the 33 to 37 stage. When regional opportunity density is low and family responsibilities rise, risk avoidance and stabilisation are strengthened. When platform competition relies heavily on visibility, title-based signal capital becomes more salient.

Implications

The findings suggest that, in practice, career planning for humanities and social sciences majors should be understood as a dynamic process of stage-based task configuration rather than as a one-off static choice aimed at locating a single optimal pathway. At the individual level, the priority sequences revealed at life course nodes make it possible to translate career planning into executable bundles of stage-based strategies. During the 18 to 22 stage, investment should focus on the verifiability and presentability of capability foundations. Specifically, this means converting abstract abilities into externally assessable outputs through internships, portfolios, and project experience, supplemented by blended learning models that incorporate targeted evaluations and improvement methods for deep learning ability [20]. Such models help ensure students develop actionable professional skills rather than superficial knowledge. Meanwhile, efforts should be made to enhance the visibility of career signals via competitions, awards, and publicly available deliverables. During the 23 to 27 stage, narratives are highly concentrated around a pathway lock in window. At this point, educational upgrading and threshold-oriented certifications are often framed as key mechanisms for entering specific tracks. Practically, however, this stage requires realistic cost benefit accounting and alignment with opportunity windows, in order to avoid misusing further education as a risk avoidance device that merely postpones decision making. During the 28 to 32 stage, narratives shift from entry to competition. A coupled strategy of intensified qualification building and project capital accumulation is therefore more appropriate, ensuring that certification and professional title preparation serve the formation and transfer of job relevant capabilities rather than devolving into inefficient over competition through

credential stacking or fragmented experience chasing. During the 33 to 37 stage, promotion and professional title pathways are clearly governed by organisational rules. Practically, evaluation and promotion should be treated as problems of documentation closure and time window management. This implies establishing institutional understanding and evidence chains in advance, thereby reducing passive depletion under uncertainty. At age 38 and above, the strategic centre of gravity is better shifted from entry competition to the reallocation of experiential capital, with experience productised through management, consulting, side ventures, or knowledge-based products. Meanwhile, selective supplementation of certifications or degrees should be adopted, targeting only those key capability gaps most likely to leverage returns and opportunities.

In parallel, employment and development support provided by universities and schools should move from generic publicity to modular programmes differentiated by year level and track. Lower year cohorts should be trained to translate capabilities into portfolio and project outputs, the mid stage should be equipped with decision support tools regarding entry regimes and qualification pathways, and senior cohorts should be supported in role targeting and signal presentation. Notably, the application of BP neural network-enhanced systems can effectively integrate employment guidance with mental health support for college students, addressing both career anxiety (a key risk narrative in this study) and psychological stress during the transition from campus to workplace. This integration in turn improves the feasibility and effectiveness of stage-based career planning [21]. Alumni pathway case repositories and industry mentor mechanisms can further reduce students' misconceptions about a single correct answer.

At the same time, practical application must retain critical awareness and boundary sensitivity. The strong emphasis on age nodes, certifications, and titles in online texts reflects real institutional constraints and competitive structures, yet it may also be amplified by platform dissemination mechanisms into anxiety driven narratives or signal worship, thereby reducing complex career development problems to consumable templates. Without careful discernment, individuals may become trapped in high cost and low return investment loops structured around window countdowns and credential stacking. Organisations may also reinforce path dependence on superficial signals in selection processes, which can in turn intensify structural mismatches for humanities and social sciences talent. Consequently, improvements at organisational and institutional levels are equally important. Firms and employing organisations should clearly distinguish between hard thresholds and developable components in recruitment, translate job requirements into verifiable task performance, portfolios, or project outcomes, reduce excessive reliance on ambiguous titles and labels, and guide employees towards transferable capabilities through staged milestones and transparent evaluation criteria. Public sector agencies and public institutions can, without lowering quality, enhance rule transparency by clearly presenting documentation requirements, cycle rhythms, and window conditions for promotion and professional title evaluation, while strengthening capability indicators aligned with humanities and social sciences strengths, such as writing and communication, policy analysis, evidence use, and public service design, thereby reducing the institutional costs associated with guessing rules. Industry associations and platform ecosystems can also build dynamically updated public knowledge bases of pathway information, standardising the presentation of threshold structures, the signalling value of certifications, typical growth pathways, and common pitfalls across tracks, thus reducing noise and misinformation. Overall, the practical value of this study lies not in delivering an absolutised optimal solution, but in enabling individuals and organisations to recognise, from a conditional optimality perspective, which stage based task configurations are more likely to yield higher career feasibility and expected returns under particular constraint and resource structures, and under what conditions identical investments may amount only to inefficient coping driven by anxiety narratives.

Conclusion

Although this study, through large scale cross platform web crawling, systematic cleaning, and NVivo assisted coding, provides a comparatively comprehensive account of the priority ordering of career planning tasks and the logic of conditional optimality across different

life course nodes among humanities and social sciences groups, several limitations should be treated with due caution. First, the corpus is primarily drawn from Zhihu and Xiaohongshu. While this choice facilitates complementary evidence from argumentative narratives and action-oriented narratives, it inevitably introduces platform level structural bias. User composition, content mechanisms, and discursive styles shape both modes of expression and issue visibility. Accordingly, the findings are closer to the career cognition structure manifested in platform based public discussion than to a strictly statistically representative picture of the entire humanities and social sciences population.

Second, the self-selection bias inherent in text-based research remains. Individuals who are more willing to articulate anxiety, redirection, successful entry experiences, or success pathways are more likely to speak, whereas the experiences of silent users and low visibility groups may be systematically underestimated. In addition, although the study improves auditability through denoising, anonymisation, and audit trails, it cannot strongly verify posters' real identities, such as disciplinary background, age, region, and organisational sector. Therefore, some information must be treated as self-reported narratives rather than fully verifiable objective facts.

Third, although age signal identification adopts explicit rules and differentiates the analytical use of texts with identifiable age signals versus those without stated ages, a substantial share of the corpus still does not allow age to be reliably extracted, with the proportion of unstated age being relatively high in the present sample. This means that evidence supporting the age band comparison matrix is more concentrated among users willing to self-disclose age, which may in turn increase the visibility of certain node narratives, such as the age 35 threshold.

Fourth, although the study controls interpretive subjectivity through dual coding, consistency assessment, theoretical saturation checks, and negative case examination, qualitative coding inevitably relies on researchers' theoretical sensitivity and boundary judgements. This is particularly salient for overlapping concepts, such as opportunity windows, age thresholds, and public sector locking in. Even with continual refinement of the codebook, interpretive contestability cannot be fully eliminated. Fifth, the study's conclusion framing in terms of conditional and contextual optimality is intended to avoid misrepresenting online experiences as universal regularities. However, this also implies that its normative implications remain strategies that are more likely to be effective under specific conditions, and cannot substitute for fine grained decision making under individual heterogeneity, industry heterogeneity, and regional institutional variation.

Future research can advance at least five directions in order to improve generalisability, explanatory power, and policy or practical value. First, data sources and comparative frameworks can be expanded. While retaining core platforms, future studies can incorporate Weibo, Bilibili, Douban, niche community forums, and recruitment platform comments to construct a more complete career discourse ecology. Systematic comparisons across platforms, user composition, and narrative mechanisms would allow platform bias to be transformed from an uncontrolled error into an explicable variable.

Second, researchers can move towards multimodal evidence and stronger evidence chains. On Xiaohongshu and short video platforms, career experiences are often constructed jointly through images, video, and comments. Future work can incorporate image and video titles, subtitles, comment interactions, and dissemination indicators to examine the mechanisms through which visibility strategies, such as titles and competitions, operate across media formats.

Third, mixed methods can be adopted for external validation. Once the text based typology and task map are preliminarily stabilised, surveys or interviews can be used to validate key propositions, such as whether the priority sequences observed by age band correspond to real behaviours, and which condition variables, including industry entry barriers, certification regimes, organisational promotion rules, regional opportunity density, and family responsibilities, significantly alter the optimal sequence. This would enable a clearer distinction between normative claims embedded in online narratives and actual patterns in real world decision making.

Fourth, the time dimension and analyses of policy or

environmental shocks can be strengthened. Although the present study covers a long time window from 2018 to 2025, future research can further refine annual and event window slices, such as labour market fluctuations, changes in examination regimes, and platform rule adjustments, in order to observe how the narrative weights attached to education, certification, and public sector choices shift with macro level environments. This would facilitate a combined explanation that integrates life course mechanisms with external shock mechanisms.

Fifth, reproducibility and methodological innovation can be enhanced. Subject to ethical compliance and anonymisation, future research can more systematically disclose crawling rules, cleaning criteria, codebook evolution, and key query settings to develop a more replicable research pipeline. In parallel, human machine collaborative coding can be explored, such as using large language models for initial screening and candidate theme generation, followed by researcher adjudication of boundaries and negative case checks, thereby maintaining interpretive quality while scaling up corpus size.

Overall, if future work can close the loop among cross platform generalisation, mixed method validation, long horizon dynamic explanation, and reproducible methodological pipelines, it will be more likely to advance the conditional optimality framework proposed in this study from a descriptive scheme into a more explanatory and operationalisable system of theoretical propositions.

Funding

This work was not supported by any funds.

Acknowledgements

The authors would like to show sincere thanks to those techniques who have contributed to this research.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Main, J. B., Prenovitz, S., Ehrenberg, R. G. (2019) In pursuit of a tenure-track faculty position: Career progression and satisfaction of humanities and social sciences doctorates. *The Review of Higher Education*, 42(4), 1309-1336.

[2] Lerro, A., Schiuma, G., Manfredi, F. (2022) Entrepreneurial development and digital transformation in creative and cultural industries: trends, opportunities and challenges. *International Journal of Entrepreneurial Behavior & Research*, 28(8), 1929-1939.

[3] Piattoeva, N. (2018) Elastic numbers: National examinations data as a technology of government. *Governing by Numbers*, 18-36.

[4] Zhang, Y. (2024) Path of career planning and employment strategy based on deep learning in the information age. *Plos One*, 19(10), e0308654.

[5] Evans, J. (2022) From Text Signals to Simulations: A Review and Complement to Text as Data by Grimmer, Roberts & Stewart (PUP 2022). *Sociological Methods & Research*, 51(4), 1868-1885.

[6] Kozinets, R. V., Gretzel, U. (2024) Netnography evolved: new contexts, scope, procedures and sensibilities. *Annals of Tourism Research*, 104, 103693.

[7] Chen, Y., Wu, X., Hu, A., He, G., Ju, G. (2021) Social prediction: a new research paradigm based on machine learning. *The Journal of Chinese Sociology*, 8(1), 15.

[8] Akkermans, J., da Motta Veiga, S. P., Hirschi, A., Marciniak, J. (2024) Career transitions across the lifespan: a review and research agenda. *Journal of Vocational Behavior*, 148, 103957.

[9] Salwén, H. (2021) Research ethical norms, guidance and the internet. *Science and Engineering Ethics*, 27(6), 67.

[10] Cui, H., Yu, P. (2026) Human-generative AI cooperation and knowledge collaboration in online knowledge communities: an analysis based on adaptive structuration theory for individuals. *Information Technology & People*, 1-38.

[11] Yu, Z. (2025) Xiaohongshu: Modernizing social commerce in China's digital landscape. *Cases on Chinese Unicorns and the Development of Startups*, 327-346.

[12] Charmaz, K., Thornberg, R. (2021) The pursuit of quality in grounded theory. *Qualitative Research in Psychology*, 18(3), 305-327.

[13] Bryda, G., Costa, A. P. (2023) Qualitative research in digital era: innovations, methodologies and

collaborations. *Social Sciences*, 12(10), 570.

[14] Olapane, E. C. (2021) An in-depth exploration on the praxis of computer-assisted qualitative data analysis software (CAQDAS). *Journal of Humanities and Social Sciences Studies*, 3(11), 57-78.

[15] Dhakal, K. (2022) NVivo. *Journal of the Medical Library Association: JMLA*, 110(2), 270.

[16] O'Connor, C., Joffe, H. (2020) Intercoder reliability in qualitative research: debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1609406919899220.

[17] Braun, V., Clarke, V. (2023) Toward good practice in thematic analysis: Avoiding common problems and be (com) ing a knowing researcher. *International Journal of Transgender Health*, 24(1), 1-6.

[18] Kitchenham, B., Madeyski, L., Budgen, D. (2022) SEGRESS: software engineering guidelines for reporting secondary studies. *IEEE Transactions on Software Engineering*, 49(3), 1273-1298.

[19] Drolet, M. J., Rose-Derouin, E., Leblanc, J. C., Ruest, M., Williams-Jones, B. (2023) Ethical issues in research: perceptions of researchers, research ethics board members and research ethics experts. *Journal of Academic Ethics*, 21(2), 269-292.

[20] Li, Y., Cheng, H., Qin, Q. (2025) Evaluations and improvement methods of deep learning ability in blended learning. *International Journal of e-Collaboration (IJeC)*, 21(1), 1-17.

[21] Deng, Z., Xiang, H., Tang, W., Cheng, H., Qin, Q. (2024) BP neural network-enhanced system for employment and mental health support for college students. *International Journal of Information and Communication Technology Education (IJICTE)*, 20(1), 1-19.