

---

Themenheft 72: Mixed Methods im Spannungsfeld. Weiterentwicklung digitaler Forschungsmethoden zwischen KI, NLP, Big Data und transdisziplinärer Forschung. Herausgegeben von Jana Heinz, Andrea Hense und Christian Janßen

## Web Scraping, Topic Modelling, and Qualitative Methods in Educational Research

### A Mixed-Methods Study on Digitalization and Educational Equity in Teacher Professional Development

Alexandra Krebs<sup>1</sup>  and Barbara Getto<sup>2</sup> 

<sup>1</sup> Katholische Universität Eichstätt-Ingolstadt

<sup>2</sup> Pädagogische Hochschule Zürich

#### Abstract

*The ongoing digital transformation poses significant new challenges for schools, education, and teacher training. The DEEP – Professionalization project, based on the IPRODiG model, investigates how digitalization and educational equity are integrated into continuing education for primary school teachers in German-speaking Switzerland. This paper presents a first-phase pilot report focusing on the systematic mapping of publicly available continuing education offerings. The objective is to identify training formats, content, objectives, and assess the extent to which these two critical topics are addressed. As an exploratory analysis of a pilot dataset comprising 4,658 professional development offerings, the paper aims to generate initial, data-grounded hypotheses rather than definitive conclusions. These hypotheses are intended as a basis for subsequent analytical phases and for future strategic development in teacher continuing education.*

*Using automated web scraping techniques, publicly available course descriptions were systematically collected from a diverse range of institutional providers. The dataset was analysed through topic modelling to identify dominant thematic structures, content priorities, and their interrelations. As these data primarily reflect the representational level of course offerings, the study adopts a mixed-methods perspective, complementing the computational analysis with questionnaires and interviews to enable triangulation and contextual interpretation. These additional data sources are announced as subsequent project phases and are not yet integrated into the present analysis.*

*The paper critically reflects on the methodological affordances and limitations of AI-supported research approaches, demonstrating how surface-level quantitative pattern detection can be productively combined with qualitative inquiry. The relevance of these methods specifically for media education research is discussed, particularly regarding digital research literacy and the epistemological positioning of computational methods within the field. The findings provide a structured overview of current trends in teacher*

*professional development and offer data-informed, preliminary insights for future strategies, particularly regarding strengthening equity-oriented professional learning in contexts of ongoing digital transformation.*

**Web Scraping, Topic Modelling und qualitative Methoden in der Bildungsforschung.  
Eine Mixed-Methods-Studie zu Digitalisierung und Bildungsgerechtigkeit in der  
Weiterbildung von Lehrpersonen**

**Zusammenfassung**

*Die fortschreitende digitale Transformation stellt Schulen, Bildung und Lehrerausbildung vor erhebliche neue Herausforderungen. Das Projekt DEEP – Professionalization, welches auf dem IPRODiG-Modell rekurriert, untersucht, wie Digitalisierung und Bildungsgerechtigkeit in die Weiterbildung von Grundschullehrkräften in der deutschsprachigen Schweiz integriert werden. Der vorliegende Beitrag versteht sich als erster Projektbericht und konzentriert sich auf die Kartierung öffentlich zugänglicher Weiterbildungsangebote als erste Projektphase. Ziel ist es, Ausbildungsformate, Inhalte und Ziele zu identifizieren und zu bewerten, inwieweit diese beiden Themen behandelt werden. Als explorative Analyse eines Pilotdatensatzes mit 4.658 Weiterbildungsangeboten zielt der Beitrag auf die Generierung erster, datengestützter Hypothesen ab, die als Grundlage für nachfolgende Analysephasen dienen sollen. Mithilfe automatisierter Web-Scraping-Techniken wurden öffentlich zugängliche Kursbeschreibungen von einer Vielzahl institutioneller Anbieter systematisch gesammelt. Der Datensatz wurde durch Topic Modelling analysiert, um dominante thematische Strukturen, inhaltliche Prävalenzen und deren Wechselbeziehungen zu identifizieren und zu untersuchen. Da diese Daten in erster Linie die repräsentative Ebene der Kursangebote widerspiegeln, verfolgt die Studie einen Mixed-Methods Ansatz, der die computergestützte Analyse durch Fragebögen und Interviews ergänzt, um eine Triangulation und kontextuelle Interpretation zu ermöglichen. Diese weiteren Datenquellen werden als nachfolgende Projektphasen angekündigt und sind im vorliegenden Beitrag noch nicht empirisch integriert. Der Beitrag reflektiert kritisch die methodischen Möglichkeiten und Grenzen von KI-gestützten Forschungsansätzen und zeigt, wie die quantitative Mustererkennung auf Oberflächenebene produktiv mit qualitativer Forschung kombiniert werden kann. Die Relevanz dieser Methoden speziell für die medienpädagogische Forschung wird diskutiert, insbesondere im Hinblick auf digitale Forschungskompetenz und die erkenntnistheoretische Verortung rechnergestützter Methoden innerhalb der Disziplin. Die Ergebnisse bieten einen strukturierten Überblick über aktuelle Trends in der beruflichen Weiterbildung von Lehrkräften und liefern datengestützte, vorläufige Erkenntnisse für zukünftige Strategien, insbesondere im Hinblick auf die Stärkung einer bildungsgerechten beruflichen Weiterbildung im Kontext der fortschreitenden digitalen Transformation.*

## 1. Introduction

The digital transformation of educational systems has emerged as one of the most pressing challenges for contemporary schooling, fundamentally reshaping teaching practices, curricular demands, and the professional competencies required of educators. As schools increasingly integrate digital technologies into their pedagogical frameworks, teachers find themselves at the intersection of two critical imperatives: mastering the affordances of digitalization while simultaneously addressing persistent issues of educational equity (Caena and Redecker 2019; Fraillon et al. 2020; Heinz 2023). This dual challenge is particularly acute in primary education, where foundational learning experiences establish trajectories that can either mitigate or exacerbate educational inequalities.

Despite widespread recognition of these challenges, teacher continuing education, a crucial lever for supporting educators in navigating digital transformation, remains insufficiently examined. While substantial research has explored initial teacher education and discrete professional development interventions (Alonzo et al. 2021; Buchner 2023), systematic analyses of the broader landscape of continuing education offerings are notably scarce (Strittmatter 1990; Hansen and Kohlstock 2022). Questions persist regarding which competencies are prioritized, how digitalization and equity concerns are integrated into professional development curricula, and whether current training formats adequately prepare teachers for the complexities of digitally mediated, equitable pedagogy. From a media education perspective, this gap is particularly salient. Investigating how digital transformation and questions of educational equity are represented in professional development offerings speaks directly to core concerns of the field: the conditions under which educators develop critical digital competencies, the institutional structures that shape digital literacy education, and the methodological potential of digital research approaches for studying these phenomena. This paper thus contributes to media educational research not only substantively, by mapping the landscape of digitalization-related professional development, but also methodologically, by demonstrating the affordances and limitations of computational methods such as web scraping and topic modelling for educational research in the digital age. However, investigating the landscape of teacher continuing education presents significant methodological challenges. How can the vast array of course offerings published on institutional websites be systematically collected, represented, and analysed? To what extent can emerging AI-supported methods such as web scraping and topic modeling be productively deployed for educational research? And what are the epistemological and practical limitations of such approaches when applied to complex educational phenomena?

This paper addresses these interconnected empirical and methodological questions through the DEEP — Professionalization project, which investigates the integration of digitalization and educational equity within continuing education for primary school teachers in German-speaking Switzerland. Grounded in the IPRoDiG model (Irion et al. 2023), the study systematically examines training formats, content priorities, and

pedagogical objectives to assess how these two domains are currently addressed in professional development offerings.

This study addresses not only scientific research questions but also explicitly pursues a Third Mission orientation: The systematic mapping of professional development offerings serves as a foundation for evidence-based recommendations to continuing education providers, educational policy, and universities. Thus, the project contributes to the societal responsibility of research institutions by generating practice-relevant knowledge that can directly inform the strategic advancement of teacher professionalization (Compagnucci and Spigarelli 2020).

Methodologically, the study employs a comprehensive mixed-methods design that critically examines the potential and constraints of automated research approaches. The present paper constitutes a first-phase report of this larger project design. In a pilot dataset of 4,658 publicly available course offerings from institutional websites across German-speaking Switzerland are systematically collected through web scraping techniques. These data are analysed using topic modeling to identify thematic clusters, key concepts, and their interrelations. Recognizing that automated methods provide primarily surface-level insights into course content, the quantitative findings will be triangulated in subsequent project phases, enabling deeper analytical understanding and contextual interpretation. In the present paper, these additional data layers are announced and described in their design logic, but are not yet empirically integrated into the analysis. The analysis provides a systematic overview of current trends and priorities in teacher professional development, yielding actionable insights for policy and practice. Initial findings generate data-based recommendations for future strategies in teacher education, particularly concerning how continuing education can more effectively support teachers in fostering equity within contexts of ongoing digital transformation. Yet equally important is the study's contribution to methodological innovation: demonstrating both the generative possibilities and necessary constraints of AI-supported research methods in capturing complex educational landscapes. Against this background, this study pursues two intertwined objectives:

Empirical Questions:

- How are digitalization and educational equity represented in the professional development landscape for primary school teachers in German-speaking Switzerland?
- Which thematic priorities and gaps can be identified?

Methodological Questions:

- How can AI-supported analytical methods (web scraping, topic modelling) be productively integrated into mixed-methods designs?
- What epistemological and practical limitations do these methods exhibit, and how can they be complemented by classical qualitative and quantitative methods?
- How does such a hybrid data integration model contribute to the methodological advancement of Mixed Methods?

## 2. Research Design

To address our research questions, we developed a mixed-methods research design (see figure 1) tailored to the complexity of the educational landscape under study. The intertwined dynamics of digital transformation and educational equity demanded an approach capable of capturing large-scale patterns while enabling nuanced interpretive insight. Accordingly, our design integrates automated data collection and computational text analysis with survey-based and qualitative inquiry. In the following sections, we outline the methodological steps of this approach, situating them in relation to comparable methodological studies. By doing so, we critically reflect on both the potential and the limitations of our chosen strategy before turning to an in-depth discussion of exemplary findings generated through this framework.

The quantitative component consists of a large-scale survey investigating teachers' digital educational equity mindset (DEEM) and their competencies in addressing the digital divide. Building on Van Dijk's (2020) Digital Divide framework, we developed and validated a measurement instrument assessing four key dimensions: fostering access to digital technologies, supporting learners' motivation, facilitating competent use, and enabling meaningful learning outcomes. The survey, administered to primary school teachers in German-speaking Switzerland, establishes baseline data on teachers' awareness and capabilities regarding digital educational equity. Validation studies demonstrated robust psychometric properties with high internal consistencies ( $\omega \geq .92$ ) and a clear four-factor structure. In the present paper, however, we focus on the second work package of the study, which centers on the systematic mapping of continuing education offerings for teachers. The survey component will be integrated in subsequent analyses.

### Inventory of Swiss Teacher Professional Development

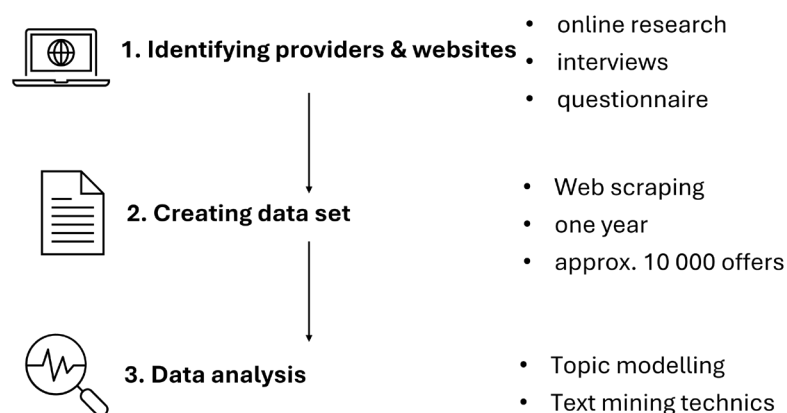


**Fig. 1:** Research Design.

As a first step (Fig. 1), we identified relevant providers and institutions offering continuing education for primary school teachers in German-speaking Switzerland. To construct a comprehensive sampling frame, we compiled an inventory of institutions and their publicly accessible websites, ensuring inclusion of both public and private providers. This process drew on multiple sources: systematic online searches, exploratory expert interviews, and questionnaire responses from key stakeholders. Through this triangulated identification strategy, we sought to capture the diversity of the professional development landscape rather than relying solely on a single information channel. For the pilot dataset, providers were selected from this inventory according to two criteria: (1) representation of both public and private continuing education providers, and (2) the availability of a sufficient number of course offerings to support the exploratory modelling and validation procedures described below. In total, 13 providers were included in the pilot dataset (Tab. 1).

| institution                                                    | public or private              | Number scraped courses |
|----------------------------------------------------------------|--------------------------------|------------------------|
| Fobizz                                                         | private                        | 198                    |
| Schule und Weiterbildung Schweiz (SWCH)                        | private                        | 121                    |
| Zürcher Arbeitsgemeinschaft<br>Lehrpersonenweiterbildung (ZAL) | private                        | 268                    |
| PH Bern                                                        | public                         | 1675                   |
| FHNW                                                           | public                         | 492                    |
| PH Wallis                                                      | public                         | 131                    |
| PH Graubünden                                                  | public                         | 112                    |
| PH Thurgau                                                     | public                         | 149                    |
| PH Luzern                                                      | public                         | 531                    |
| PH Zug                                                         | public                         | 170                    |
| PH Schaffhausen                                                | public                         | 68                     |
| PH St. Gallen                                                  | public                         | 261                    |
| PH Zürich                                                      | public                         | 482                    |
| <b>13 institutions</b>                                         | <b>3 private<br/>10 public</b> | <b>4658 courses</b>    |

**Tab. 1:** Pilot dataset.

From these institutional websites, the publicly available course information was systematically collected over a two-month period between May and June 2025. Some providers contributed substantially more data than others. PH Bern alone accounted

for 36 % of all courses, while the four largest providers together accounted for 68 % of the dataset. We did not adjust for this imbalance during the modelling process, as the study explicitly aimed to map the existing landscape of continuing education offerings as it is institutionally represented. From this perspective, weighting providers according to their actual presence and volume within the field was methodologically appropriate and analytically consistent with the exploratory character of the study.

A related methodological decision concerned how to handle courses that were offered repeatedly in similar or identical form. Each offering therefore received a unique ID. We did not examine whether two IDs referred to the same course described using highly similar wording. If a provider listed the same course across multiple semesters or at different campuses, all instances remained within the dataset. This decision was intentional: each ID represented one actual teaching offering. Consequently, repeated instances of the same course were treated not as duplicates, but as distinct professional development events within the continuing education landscape.

Following these sampling and data management decisions, and the fact that a substantial proportion of the professional development platforms examined employ dynamic interfaces (Fig. 2), a robust scraping strategy capable of handling heterogeneous, evolving data structures was selected. The automated collection of course information was implemented through Python-based web scraping procedures using the Selenium WebDriver library (Selenium, 2025). To ensure methodological rigour, the data collection process was implemented as a staged procedure that addressed legal compliance, technical execution, data processing, and quality assurance.



Fig. 2: Screenshot dynamic website interface.

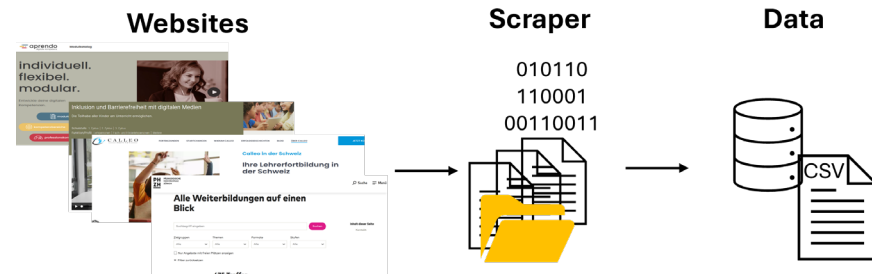


**Fig. 3:** Web scraping process.

First, particular attention was devoted to legal and ethical compliance (Kazmali and Sayar 2025). Prior to data extraction, all targeted websites were reviewed for their terms of use, copyright statements, and technical access restrictions. The scraping procedures were designed to comply with robots.txt directives and operate within permitted usage boundaries. This precaution was essential given the legal sensitivities associated with harvesting publicly available digital content for research purposes.

The second phase involved implementing the automated data collection process (Fig. 3). All scraping routines were developed in Python (version 3.x) (Mitchell 2024), incorporating key libraries for browser automation and data handling, including Selenium and Pandas. The WebDriver was configured to support dynamic page interactions and adapted to the requirements of different browser environments. Data extraction followed a structured, multi-step strategy. First, the automated environment was initialized by configuring WebDriver, applying wait routines to prevent race conditions in which scraping scripts attempt to access web elements before they are fully rendered, and integrating exception-handling protocols. Subsequently, target variables were defined. Extraction focused on key informational components relevant to course content, including course titles, subtitles, descriptions, thematic content, learning objectives, and formats. To retrieve these data reliably across heterogeneous page structures, we employed both CSS selectors and XPath expressions. Missing or inconsistently formatted entries were tracked through error-handling routines to minimize data loss and ensure reproducibility.

A major challenge in this process was the substantial heterogeneity in website design and structure across providers. Because the layout and underlying HTML architecture varied considerably, selectors had to be identified and adapted individually for each website. Accordingly, we conducted a systematic inspection and analysis of the HTML code to identify relevant elements and ensure accurate extraction.

Once collected, all data were processed in a structured manner. Extracted text fields were cleaned and reformatted into a consistent schema. Validation routines were implemented to detect incomplete or duplicate entries, and the resulting dataset was stored in a structured format suitable for subsequent computational analysis, primarily as comma-separated value (CSV) files (Tab. 2).

| title | description | goals | contents | format | target audience | other |
|-------|-------------|-------|----------|--------|-----------------|-------|
|-------|-------------|-------|----------|--------|-----------------|-------|

**Tab. 2:** Extracted text fields.

Quality assurance measures were in place throughout all stages of the scraping pipeline. From a technical perspective, automated delays were integrated between requests to minimize server load and reduce the risk of incomplete data capture. Sample-based checks were conducted to verify extraction accuracy and assess the integrity of the stored data.

Web scraping has become well established in specific research fields, particularly marketing. For example, Debón and Domenech (2024) analysed the websites of 115 wineries in the Valencia region of Spain to assess their «digital footprint» and related competitiveness. The researchers used web scraping to extract all relevant information from these sites automatically, then evaluated the data using machine learning algorithms across different clusters. Compared to manual methods, web scraping offers a more efficient way to collect data for such analyses. Additionally, this data is currently available for free public access, and particularly suitable for providing an overview of a research area or for tracking changes and developments over time. Similarly, Ashouri et al. (2024) collected data from 60,942 company websites to examine their levels and measures of digitalization. However, they also highlighted the challenges associated with this type of data collection: «Despite these advantages, challenges remain, such as potential biases in data, issues with accuracy, [...] and variations in language and website quality» (Ashouri et al. 2024, 4). It is therefore important to consider which research questions can be addressed using the available data and to acknowledge its limitations. In the present study of Swiss continuing education programs, online information generally provides only brief descriptions, particularly of course content and learning objectives. It does not fully capture the topics covered or indicate which courses teachers attend. The website data should therefore be complemented by other sources, including interviews and questionnaires conducted as part of the project. In accordance with the principles of Open Science and Open Research Data, all project data will be made available through an institutional repository after project completion. The dataset includes:

- The scraped course descriptions (anonymized according to data protection guidelines)
- The complete Python code for web scraping and topic modelling (GitHub repository)
- Topic modelling parameters and results for replication
- Documentation of preprocessing steps

This transparency enables other researchers to ensure reproducibility of the analyses and contributes to the methodological advancement of digital research methods in educational sciences (Boyd and Crawford 2012).

### 3. Data Analysis

Having established a structured corpus of course descriptions, the next phase of the study focused on analyzing the textual content to identify different themes and their relationships to digitalization and educational equity. Because automated scraping alone produces descriptive data, subsequent analytical steps were required to uncover latent conceptual patterns that could inform substantive interpretation.

To address this analytical challenge, methods were required that move beyond descriptive aggregation and enable the systematic detection of latent thematic structures within large text corpora. Topic modelling offers such an approach by providing computational techniques for identifying recurring semantic patterns and organizing them into interpretable thematic clusters. The following section, therefore, situates topic modelling as both a methodological and analytical foundation for examining the scraped course descriptions before detailing its specific implementation in this study.

Topic modelling has become a widely adopted analytical approach for exploring large-scale textual datasets across diverse research domains. As Churchill and Singh (2022) note, «topic models have been applied to everything from books to newspapers to social media posts in an effort to identify the most prevalent themes of a text corpus» (p. 1). In their comprehensive review, the authors trace the development of topic modelling from its early statistical foundations in the 1990s to contemporary neural approaches, highlighting its growing relevance for interdisciplinary research. At its core, topic modelling seeks to identify latent thematic structures by detecting recurring word patterns and grouping them into topics that reflect the underlying content of documents (Churchill and Singh 2022). Owing to this capacity, topic modelling has been employed in fields ranging from historical and literary studies to climate science and political communication. For instance, researchers have analysed historical newspapers to uncover long-term thematic trends (Newman and Block 2006), examined the evolution of climate change discourse (Sleeman et al. 2017), identified dominant themes in nineteenth-century literature (Jockers and Mimno 2013), and explored public debates on social media platforms such as Twitter, including parenting discussions (Ryan et al. 2021) and electoral campaigns (Bode et al. 2020). Therefore, topic modelling offers a well-established yet evolving methodological lens for analyzing large corpora of professional development descriptions and for systematically exploring how themes of digitalization and educational equity are articulated within continuing education for teachers.

The study's analytical framework comprised multiple stages. Topic modelling was selected as the primary computational technique due to its ability to identify hidden thematic structures in large text datasets. We reviewed different modelling approaches commonly used in natural language processing and educational research, such as Latent Dirichlet Allocation (LDA), Non-negative Matrix Factorization (NMF), Latent Semantic Analysis (LSA), and clustering-based methods, recognizing that each offers distinct

advantages and limitations depending on dataset characteristics and research purposes. For the present study, BERTopic (Grootendorst 2022) was chosen due to its capacity to handle heterogeneous textual content without requiring a pre-specified number of topics. Its combination of transformer-based embeddings and hierarchical clustering enables greater semantic sensitivity and flexibility than more traditional probabilistic approaches, albeit at the cost of computational complexity.

The objective of our topic modelling process was to identify dominant areas of professional development provision, with particular emphasis on how digitalization and educational equity are represented across the corpus. All analyses were conducted in Python using relevant libraries, including BERTopic, transformers, pandas, and numpy.

Before modelling, the dataset underwent a series of preprocessing steps intended to standardize the textual data and enhance the interpretability of the resulting topics. These steps included cleaning the text, reducing words to their base forms, removing highly frequent stop words, and preserving meaningful multi-word expressions: First, we fixed line breaks and extra spaces. Next, we used a German language tool called spaCy to find the base form of each word. For example, the words *lerne* and *lernten* both became *lernen*. We then dropped very short words and words that were too common to be useful — words like *und*, *der*, or *wir*. The list contained 1,121 such words. We kept single words and pairs of words (for example, *digitale Medien*). The result was a clean list of base words for each course.

Two analytical foci guided the interpretation of topics: digitalization and educational equity. The former encompassed issues such as digital tools, learning platforms, media education, and professional digital competencies. The latter included differentiation, individualized support structures, inclusion, and broader questions of opportunity and participation. Comparative inspection of topic clusters enabled us to examine intersections and divergences between these domains.

The BERTopic analysis identified twelve distinct topics within the pilot dataset (see Table 3). To derive these topics, the model detected recurring linguistic patterns across the corpus and grouped them into higher-order thematic clusters that reflect the substantive content of the documents, in this case, professional development offerings.

We checked the quality of our topics in four ways. First, for NMF we tested every number of topics from 3 to 20. We measured how large the topics were, how much they overlapped, and how many different words they used. Second, we ran a very different second method (BERTopic) and checked whether the same themes appeared. If two very different methods produced similar results, we considered the findings more robust. Third, we examined each provider individually. A meaningful theme should appear across multiple providers, whereas a theme that appeared only within a single provider was more likely to reflect provider-specific wording or formatting practices. Fourth, we read real example courses. For every topic, we selected the three most typical courses and analysed them in full.

Each topic is represented by a set of the ten most salient terms, which illustrate the underlying semantic patterns identified by the model. The accompanying topic labels constitute interpretive summaries assigned by the researchers and are not generated by the model itself. As such, the analytical outcome reflects an interplay between computational pattern detection and human sense-making. The coherence of the extracted term clusters suggests that the identified topics are semantically meaningful and can be interpreted in a significant way. Thus, the topic modeling process allows us to identify the most common themes within the text corpus.

|                | <i>label</i>                       | <i>words</i>                                                                                                                                                                                                                          |
|----------------|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>topic 0</b> | Learning and Teaching              | learn, school, teacher, familiarise, situation, child, develop, teaching, reflect, exchange<br>(lernen, schule, lehrperson, kennen, situation, kind, entwickeln, unterricht, reflektieren, austausch)                                 |
| <b>topic 1</b> | Language Education                 | teaching, German, familiarise, learn, language, language support, teaching materials, language, child, second language<br>(unterricht, deutsch, kennen, lernen, sprachlich, sprachförderung, lehrmittel, sprache, kind, zweitsprache) |
| <b>topic 2</b> | Arts Education                     | creative, image, teaching, learning, technique, different, creativity, design, designing, material<br>(kreativ, bild, unterricht, lernen, technik, verschieden, kreativität, gestalt, gestalten, material)                            |
| <b>topic 3</b> | Digital Media                      | digital, medium, teaching, familiarise, tools, learning, different, computer science, tablet, create<br>(digital, medium, unterricht, kennen, tools, lernen, verschieden, informatik, tablet, erstellen)                              |
| <b>topic 4</b> | Sports and Games                   | play, teaching, physical education, child, movement, learning, familiarise, game, foster, practical<br>(spiel, unterricht, sportunterricht, kind, bewegung, lernen, kennen, game, fördern, praktisch)                                 |
| <b>topic 5</b> | Music Education                    | musical, music teaching, music, song, voice, lesson, familiarise, play music, singing, learning<br>(musikalisch, musikunterricht, musik, lied, stimme, unterricht, kennen, musizieren, song, lernen)                                  |
| <b>topic 6</b> | Nature and Environmental Education | nature, outdoors, learning, visit, teaching, plant, opportunity, fire, climate change, animal<br>(natur, draussen, lernen, besuch, unterricht, pflanze, möglichkeit, feuer, klimawandel, tier)                                        |
| <b>topic 7</b> | Individual Support                 | mas, integrative, if, mas if, competence, coaching, degree program, school, support, learning<br>(mas, integrativ, if, mas if, kompetenz, coaching, studiengang, schule, förderung, lernen)                                           |

|                 | <i>label</i>   | <i>words</i>                                                                                                                                                                                                                                                                  |
|-----------------|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>topic 8</b>  | Mathematics    | mathematics, mathematical, math book, math lessons, learning, digital, teaching materials, exercises, individual, new edition math book<br>(mathematik, mathematisch, mathbuch, mathematikunterricht, lernen, digital, lehrmittel, aufgabe, individuell, neuausgabe mathbuch) |
| <b>topic 9</b>  | AI             | ai, intelligence, artificial, artificial intelligence, chatGPT, teaching, artificial intelligence, artificial, use, intelligence ai<br>(ki, intelligenz, künstlich, künstlich intelligenz, chatgpt, unterricht, künstliche intelligenz, künstliche, einsatz, intelligenz ki)  |
| <b>topic 10</b> | Neurodiversity | autism, adhd, ad, asd, child, understanding, autism AD, human, learning, behavior<br>(autismus, adhs, ad, ass, kind, verstehen, autismus ad, mensch, lernen, verhalten)                                                                                                       |
| <b>topic 11</b> | Diversity      | diversity, racism, school, approach, anti-racism, teaching, teaching materials, thinking, tamasha<br>(vielfalt, rassismus, schule, umgang, diversität, rassismuskritisch, unterricht, lehrmittel, denken, tamasha)                                                            |

**Tab. 3:** Topics labels and words.

An examination of the identified topics reveals a structured composition of the professional development landscape captured in the dataset. On the one hand, the corpus comprises broadly pedagogical offerings that place teachers, learners, and instructional situations at the center (Topic 0). On the other hand, a substantial share of the topics reflects subject-specific professional development aligned with core areas of the primary school curriculum, including language education (Topic 1), arts education (Topic 2), physical education (Topic 4), mathematics (Topic 8), environmental and nature-based learning (Topic 6), and music education (Topic 5). Given the data's institutional context, this distribution is unsurprising, as it mirrors both the central actors in classroom instruction and the foundational subject domains of primary schooling.

Beyond these expected structures, the topic model also identifies clusters that explicitly foreground cross-cutting educational challenges. These include offerings related to digitalization, such as digital media and educational technologies (Topic 3) as well as artificial intelligence in education (Topic 9), alongside topics that address dimensions of educational equity, including individualized support and inclusive practices (Topic 7), neurodiversity and special educational needs (Topic 10), and social diversity and anti-discrimination (Topic 11).

From an interpretive perspective, this pattern may tentatively suggest a layered structure within the continuing education landscape. Traditional subject-oriented and general pedagogical offerings constitute a foundational layer, while digitalization and equity-oriented topics form a complementary, yet more specialized, layer. It should

be emphasized, however, that this interpretation is based on the publicly visible, communicative representation of offerings in a pilot dataset; it does not yet reflect enacted pedagogical practice or the full institutional landscape. One plausible, preliminary hypothesis – to be examined in subsequent analytical phases – is that issues of digital transformation and educational equity are not yet systematically embedded across subject-specific professional development but are instead predominantly addressed in dedicated, thematically isolated formats. This structural separation may reflect institutional logics of continuing education provision, in which emerging cross-cutting challenges are treated as additional competencies rather than as integral dimensions of everyday teaching practice.

Figure 4 illustrates the proportional distribution of topics across the dataset. The visualization indicates a clear predominance of general pedagogical offerings, which account for 41 % of all identified topics, alongside subject-specific professional development formats, which together comprise 39 % of the corpus. Within the latter category, language education accounts for the largest share (13 %). In contrast, topics explicitly addressing digitalization represent a comparatively small share of the offerings (14 %), while equity-related themes account for only a marginal 7 %.

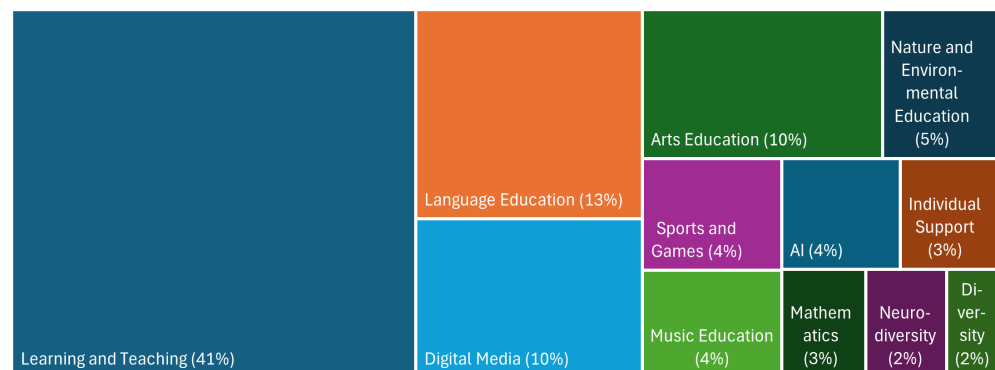


**Fig. 4:** Topics distribution.

Taken together, these distributions warrant cautious interpretation. The pilot dataset is dominated by a small number of large providers (notably PH Bern, 36 %), and the proportions therefore reflect institutionally weighted representation rather than an even landscape. Beyond confirming thematic imbalances, this distribution indicates differentiated levels of curricular stabilization across the professional development landscape. As a tentative hypothesis, general pedagogical and subject-specific topics appear to benefit from long-standing curricular normalization, that is, their stable integration into institutional curricula, professional standards, and routine training formats. In contrast, digitalization and equity-related themes seem more weakly institutionalized. As a result, offerings in these latter domains may be more dependent on short-term

policy initiatives, project-based funding, or individual provider expertise, leading to lower overall prevalence but potentially greater thematic volatility. These hypotheses will require validation through the expanded dataset and through triangulation with survey and interview data in subsequent project phases.

The hypotheses derived from the topic modelling can now be examined, at least exemplarily, through a qualitative inspection of selected course descriptions. The excerpts presented below were selected as prototypical examples of their respective topics. More specifically, they were chosen because their content closely reflected the dominant keywords and semantic patterns identified by the topic model and thus represented particularly characteristic instances of the corresponding thematic cluster. In this sense, the examples can be understood as cases located close to the semantic centre of each topic. Their purpose is not to provide a representative qualitative sample of all courses within a topic, but rather to explore how the computationally identified themes are reflected in actual course descriptions and to assess the extent to which the resulting topic clusters correspond to meaningful pedagogical distinctions. For reasons of analytical focus and comparability, the analysis concentrates exclusively on course titles and descriptions, or selected parts thereof.

A first qualitative comparison reveals a clear distinction between general pedagogical offerings (Topic 0) and subject-specific professional development formats (here exemplified by Topics 1, 2, 5, and 6). General pedagogical courses primarily address communicative, reflective, and relational dimensions of teaching and learning:

Topic 0 (Learning and Teaching): «More cooperation through better conversations. Discover new strategies for challenging conversations. In this course, you will learn how to reduce resistance, conduct direct and appreciative conversations, and remain confident» (fhnw\_inhalte-0001).

In contrast, subject-specific topics are clearly anchored in particular curricular domains and emphasize concrete instructional practices, materials, and learning activities. This becomes evident in examples from language education (Topic 1), environmental education (Topic 6), music education (Topic 5), and creative arts education (Topic 2):

Topic 1 (Language Education): «Discovering basic handwriting through scribble stories. Listening to a story, scribbling along, and finally searching for letters in the scribbles like a detective – an introduction to the world of letters that is not only fun, but also playfully develops fine motor skills and basic writing forms.» (fhnw\_inhalte-0051)

Topic 6 (Nature and Environmental Education): «Exploring and experimenting at bodies of water. What can be discovered at streams, ponds, and springs? How do students in cycles 1 and 2 explore and document these habitats? Tested teaching

ideas lead to diverse, competence-oriented learning at bodies of water» (fhnw\_inhalte-0021).

Topic 5 (Music Education): «Music with the class. An introductory course. Ideas for classroom music-making with Orff instruments. Learning how to accompany selected songs with instruments and discovering ways to use age-appropriate instruments in diverse ways in the classroom» (ph\_bern\_tm2-0028).

Topic 2 (Arts Education): «Fascinating surface designs. Inspired by the aesthetics and uniqueness of the walls of the Attisholz site, we create exciting textured images using construction materials such as concrete and plaster, as well as colors and charcoal» (fhnw\_inhalte-0167).

A similarly clear separation emerges between topics addressing educational equity (Topics 7, 10, and 11) and those focusing on digitalization (Topics 3 and 9). Equity-oriented offerings explicitly foreground inclusion, learner diversity, and differentiated support structures, as illustrated by the following examples:

Topic 7 (Individual Support): «CAS Diagnostic Support and Learning Guidance. You work with students with learning difficulties and wish to deepen your professional competencies. You are interested in becoming a special education teacher and would like to expand your knowledge through continuing education as a first step» (fhnw\_inhalte-0082).

Topic 10 (Neurodiversity): «Autism in school. You already have basic knowledge of autism spectrum disorders (ASD) but are unsure how to translate this knowledge into pedagogical practice within the classroom team and classroom setting. This course addresses these questions in relation to your current professional context» (fhnw\_inhalte-0041).

Topic 11 (Diversity): «Transgender in school. This course for school leaders provides a comprehensive insight into the needs and challenges of trans students. You will be empowered to create an inclusive and supportive environment for your trans students» (fhnw\_inhalte-0442).

In contrast, digitalization-oriented topics emphasize technological tools, platforms, and emerging innovations in educational contexts.

Topic 3 (Digital Media): «Creating digital (picture) books. Creating e-books and digital picture books. The module 'Creating digital (picture) books' is primarily aimed at teachers of cycles 1 and 2. In this module, you will engage intensively with the e-book tool «Bookcreator»» (aprendo\_inhalte-0006).

Topic 9 (Artificial Intelligence): «Fundamentals of AI. A short introduction. The topic of artificial intelligence (AI) has gained significant relevance in recent months. As a teacher, it is crucial to be informed about technology and its impact on society, education, and the world of work. This module focuses on a basic

understanding of AI and AI-based tools such as ChatGPT in educational contexts in order to better assess their potential and limitations» (aprendo\_inhalte-0043).

The example concerning artificial intelligence is particularly relevant, as it highlights the rapid pace of technological change by stating that «artificial intelligence (AI) has gained significant relevance in recent months.» This supports the tentative hypothesis that such topics are closely tied to quick innovation cycles, making them less likely to be firmly integrated into long-term curricular structures.

At the same time, the qualitative inspection reveals that equity-related concerns are not exclusively confined to explicitly labelled equity topics. Subject-specific offerings also address such issues, as illustrated by mathematics-related professional development focusing on learning difficulties:

Topic 8 (Mathematics): «Dyscalculia? Arithmetic weakness? Learning difficulties! What is dyscalculia? How do arithmetic difficulties arise? Where and why do learning difficulties in mathematics frequently develop? These and other questions are discussed and answered from a mathematics education perspective» (fhnw\_inhalte-0136).

Finally, some offerings categorized under sports and play extend beyond traditional physical education and foreground game-based pedagogical principles:

Topic 4 (Sports and Games): «Games and gamification in the classroom – basic. In what ways can games enrich teaching? You will explore a variety of practical examples, try out games directly, and develop ideas for your own teaching» (fhnw\_inhalte-0189).

This preliminary observation suggests that equity concerns may be more dispersed across the offering landscape than the topic model's categorical structure implies. At the same time, the qualitative examples both support and qualify the topic modelling results. While they confirm that the identified topics reflect meaningful pedagogical patterns, they also reveal overlaps and cross-cutting concerns that are not fully captured by discrete topic categories. This demonstrates the value of combining automated topic modelling with qualitative interpretation to capture both dominant structures and contextual nuances.

#### **4. Methodological Reflection: Mixed Methods in the Digital Age**

This study contributes to the methodological advancement of Mixed Methods in the context of digital research innovations (Ivankova 2015; Palinkas et al. 2019). While classical mixed-methods designs typically combine qualitative and quantitative methods of empirical social research (e.g., interviews + surveys), our approach integrates

Computational Social Science (CSS) methods as an independent data layer. Web scraping and topic modelling generate digital behavioral data (in the sense of: What is publicly communicated?), which differs from both self-reported data (surveys) and narratives (interviews). This extends triangulation by a third, data-driven perspective. For media education research specifically, this methodological approach opens two important avenues. First, it offers a scalable tool for systematically mapping the institutional landscape of digitalization-related teacher professional development – a domain of direct relevance to the field. Second, it raises epistemological questions about the use of digital data and AI-supported methods in educational research that are themselves central to media pedagogical inquiry: questions about data literacy, algorithmic transparency, and the conditions under which digital trace data can be considered meaningful evidence of educational reality.

Our research design combines:

- Big Data analyses (n=4,658+ course descriptions) for identifying macrostructural patterns
- Survey data for capturing individual competencies and needs
- Qualitative interviews for contextualization and deepening

This integration corresponds to hybrid data integration models (Deterding and Waters 2021; Kuckartz and Rädiker 2023), the systematic linkage of Big Data analyses with classical methods to map macrosocial structures (here: professional development landscape) and microdynamic processes (here: teacher professionalization trajectories) in an integrated analytical approach. In the present paper, only the first layer – the computational mapping – is empirically realized; the remaining layers constitute the design logic for subsequent phases.

A central result of our methodological reflection concerns the epistemic limits of CSS methods: Topic modelling can capture representation levels (what is communicated and how?) but not practice levels (what happens in professional development courses?). This insight underscores that AI-supported methods do not replace classical mixed-methods approaches but complement them with an additional data layer, and with specific epistemic potentials and limitations (Elish and Boyd 2018; Salganik 2019).

## **5. Discussion and Limitations**

While computational text analysis provides valuable insights into the thematic orientations of professional development offerings, important methodological limitations must be acknowledged. A central concern relates to the epistemic status of scraped data. Course descriptions published on institutional websites primarily convey the intended positioning, framing, or marketing of continuing education programs rather than their enacted pedagogical reality. Consequently, topic modelling captures predominantly

surface-level discursive patterns, revealing how providers present and structure their offerings, but not how learning processes unfold in practice or which competencies teachers ultimately develop. This distinction is particularly salient for research questions addressing digitalization and educational equity, as espoused intentions may diverge substantially from experienced professional learning.

The topic modelling results presented in this paper should be understood as exploratory mappings of publicly visible representations of professional teacher development, constituting tentative, data-grounded hypotheses that require validation through broader datasets and complementary methods. The identified topics constitute probabilistic thematic clusters that require cautious interpretation and contextualization, especially when normatively charged concepts such as equity are involved.

Importantly, the present analysis is based on a pilot dataset comprising 4,658 professional development offerings. In a subsequent project phase, the analytical procedures described in this paper will be extended to the full dataset of more than 10,000 scraped course offerings. This expanded analysis will allow for systematic comparison between the pilot and full datasets and will enable assessment of whether the thematic structures, distributions, and interpretive patterns identified here remain stable, shift, or reveal additional dimensions when applied at a larger scale.

To address the limitations of representational data, the DEEP-Professionalization project is designed as a mixed-methods research in which automated text analysis represents only one analytical component. In subsequent project phases, the computational findings will be systematically supplemented with interpretive approaches. A questionnaire component will provide quantitative data on teachers' self-assessed competencies across key dimensions of the digital divide, enabling the identification of competence profiles and professional development needs. Initial validation studies of the instrument with teachers from the United States ( $N = 327$ ) and the United Kingdom ( $N = 225$ ) have demonstrated satisfactory validity and reliability, establishing a robust basis for future data collection in the Swiss context.

Semi-structured interviews will further extend this quantitative layer by offering contextualized insights into teachers' reasoning about educational inequality, their strategies for addressing diverse learner needs, and the institutional constraints they encounter in everyday practice. Through thematic analysis, these interviews will allow for a deeper understanding of how digital educational equity is interpreted and enacted at the classroom level dimensions that remain inaccessible through website-based analyses alone.

In addition, focus group discussions with continuing education providers will complement the teacher-centered perspective by examining the supply side of professional development. These discussions will explore how institutions conceptualize equity and digitalization, how programmatic decisions are made, and which structural conditions shape the design of professional development offerings. Case-based and material

analyses of selected programs will further support this triangulation by enabling closer inspection of pedagogical depth, instructional design, and curricular integration.

Overall, this triangulation strategy positions computational topic modelling not as a stand-alone explanatory device, but as one element within a broader methodological framework. By combining large-scale pattern detection with qualitative and quantitative inquiry, the project aims to enhance analytical robustness and to bridge the gap between mapped representations of professional development and the lived realities of teacher professionalization.

The findings of this study aim to inform multiple stakeholders: (1) Continuing education providers can use the identified thematic gaps to strategically develop their offerings; (2) Educational policy actors receive data-based insights for prioritizing funding focuses; (3) Teachers can make informed decisions about their professional development. This transdisciplinary orientation among research, educational practice, and policy aligns with growing expectations for socially relevant research.

A further contribution to methodological transparency lies in the planned publication of all research data and analysis scripts. This not only enables the reproducibility of our results but also provides resources for methodological training and comparative studies in other contexts. For the field of media education, the study additionally provides a methodological template for how web-based institutional data can be systematically analysed to track the presence – and absence – of digitalization-related themes in professional learning ecosystems.

## **6. Mapping the Gap**

This paper explored how digitalization and educational equity are represented within continuing education offerings for primary school teachers in German-speaking Switzerland. Based on an exploratory analysis of a pilot dataset comprising 4,658 publicly available course descriptions, AI-supported topic modelling was used to identify dominant thematic structures and to examine the relative positioning of equity- and digitalization-related topics within the professional development landscape. The paper constitutes a first project report, presenting exploratory findings and tentative hypotheses that will be further examined and validated in subsequent analytical phases.

The findings tentatively reveal a clear structural imbalance: general pedagogical and subject-specific offerings form the core of continuing education, while digitalization, particularly artificial intelligence, and educational equity appear less prevalent and are often addressed in isolated formats. At the same time, the results reflect publicly visible representations of professional development rather than enacted practice and must therefore be interpreted with caution.

Looking ahead, the analytical framework will be extended to the full dataset of more than 10,000 offerings to examine the stability of these patterns at scale. Subsequent project phases will triangulate the mapping results with survey and interview data to relate institutional representations to teachers' actual professional practices and professionalization trajectories. From a media pedagogical perspective, the study demonstrates that computational methods offer a promising complement to classical research designs for investigating the institutional conditions of digital education – while simultaneously underlining the indispensability of qualitative approaches for meaningful interpretation. In this sense, mapping the gap serves as a starting point for critically examining how digitalization and educational equity can be more sustainably embedded in teacher professionalization.

## References

- Alonzo, Dennis, Ria Asih, Cherry Zin Oo, Giovanni Pelobillo, and Rex Lim. 2024. «Key Elements of Effective Mandatory Professional Development Programs». *International Journal of Instruction* 17 (3): 355–370 <https://doi.org/10.29333/iji.2024.17319a>.
- Ashouri, Sajad, Arash Hajikhani, Arho Suominen, Lukas Pukelis, and Scott W. Cunningham. 2024. «Measuring digitalization at scale using web scraped data». *Technological Forecasting and Social Change* 207:123618. <https://doi.org/10.1016/j.techfore.2024.123618>.
- Bode, Leticia, Ceren Budak, Jonathan Ladd, Frank Newport, Josh Pasek, Lisah Sing, Stuart Soroka, and Michael Traugott. 2020. *Words That Matter: How the News and Social Media Shaped the 2016 Presidential Campaign*. Washington, DC: Brookings Institution.
- Boyd, Danah, and Kate Crawford. 2012. «Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon». *Information, Communication & Society* 15 (5): 662–79. <https://doi.org/10.1080/1369118X.2012.678878>.
- Buchner, Josef. 2023. «Effekte eines Augmented Reality Escape Games auf das Lernen über Fake News». *MedienPädagogik* 51:65–86. <https://doi.org/10.21240/mpaed/51/2023.01.12.X>.
- Caena, Francesca, and Christine Redecker. 2019. «Aligning teacher competence frameworks to 21<sup>st</sup> century challenges: The case for the European Digital Competence Framework for Educators (Digcompedu)». *Euro J of Education* 54 (3): 356–69. <https://doi.org/10.1111/ejed.12345>.
- Churchill, Rob, and Lisa Singh. 2022. «The Evolution of Topic Modeling». *ACM Comput. Surv.* 54 (10s): 1–35. <https://doi.org/10.1145/3507900>.
- Compagnucci, Lorenzo, and Francesca Spigarelli. 2020. «The Third Mission of the university: A systematic literature review on potentials and constraints». *Technological Forecasting and Social Change* 161:120284. <https://doi.org/10.1016/j.techfore.2020.120284>.
- Debón, Ana, and Josep Domenech. 2024. «Digital footprint approach for the study of competitiveness in wineries». *Expert Systems with Applications* 257:1–15. <https://doi.org/10.1016/j.eswa.2024.125049>.

- Deterding, Nicole M., and Mary C. Waters. 2021. «Flexible Coding of In-depth Interviews: A Twenty-first-century Approach». *Sociological Methods & Research* 50 (2): 708–39. <https://doi.org/10.1177/0049124118799377>.
- Elish, Madeleine C., and Danah Boyd. 2018. «Situating methods in the magic of Big Data and AI». *Communication Monographs* 85 (1): 57–80. <https://doi.org/10.1080/03637751.2017.1375130>.
- Julian, John Ainley, Wolfram Schulz, Tim Friedman, and Daniel Duckworth. 2020. *Preparing for Life in a Digital World*. Cham: Springer International Publishing. Cham: Springer International Publishing.
- Grootendorst, Maarten. 2022. «BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure». <https://arxiv.org/pdf/2203.05794>.
- Digitalisierung in der Weiterbildung: Umfrage zum Stand Juni 2021». [https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Kammern/Kammer\\_PH/Dokumente\\_Berichte/220525\\_Bericht\\_Digitalisierung\\_WB.pdf](https://www.swissuniversities.ch/fileadmin/swissuniversities/Dokumente/Kammern/Kammer_PH/Dokumente_Berichte/220525_Bericht_Digitalisierung_WB.pdf).
- Irion, Thomas, Traugott Böttinger, and Rudolf Kammerl. 2023. *Professionalisierung für Digitale Bildung im Grundschulalter. Ergebnisse des Forschungsprojekts P<sup>3</sup>DiG*. Münster: Waxmann Verlag GmbH. <https://doi.org/10.31244/9783830996415>.
- Nataliya V. 2015. *Mixed Methods Applications in Action Research: From Methods to Community Action*. Thousand Oaks California: SAGE Publications, Inc. <https://doi.org/10.4135/9781071909843>.
- Jockers, Matthew L., and David Mimno. 2013. «Significant themes in 19th-century literature». *Poetics* 41 (6): 750–69. <https://doi.org/10.1016/j.poetic.2013.08.005>.
- Kazmali, Ahmet Sinan, and Ahmet Sayar. 2025. «Web Scraping: Legal and Ethical Considerations in General and Local Context – A Review». *Procedia Computer Science* 259:1563–72. <https://doi.org/10.1016/j.procs.2025.04.111>.
- Kuckartz, Udo, and Stefan Rädiker. 2023. «Using Software for Innovative Integration in Mixed Methods Research: Joint Displays, Insights and Inferences with MAXQDA». In *The Sage Handbook of Mixed Methods Research Design*, edited by Cheryl Poth, 315–27. London: SAGE. <https://doi.org/10.4135/9781529682663.n32>.
- Mitchell, Ryan. 2024. *Web Scraping with Python: Data Extraction from the Modern Web*. 3<sup>rd</sup> ed. Sebastopol, CA: O'Reilly Media.
- Newman, David J., and Sharon Block. 2006. «Probabilistic topic decomposition of an eighteenth-century American newspaper». *Journal of the American Society for Information Science and Technology* 57 (6): 753–67. <https://doi.org/10.1002/asi.20342>.
- Palinkas, Lawrence A., Sapna J. Mendon, and Alison B. Hamilton. 2019. «Innovations in Mixed Methods Evaluations». *Annual review of public health* 40:423–42. <https://doi.org/10.1146/annurev-publhealth-040218-044215>.
- Ryan, Rebecca, Pamela Davis-Kean, Leticia Bode, Juliane Krüger, Zeina Mneimneh, and Lisa Singh. 2021. *Parenting Online: Analyzing Information Provided by Parenting-Focused Twitter Accounts*. <https://doi.org/10.31234/osf.io/n3teh>.
- Salganik, Matthew J. 2019. *Bit by Bit: Social Research in the Digital Age*. First paperback printing. Princeton Oxford: Princeton University Press.

- Selenium. 2025. <https://www.selenium.dev/>. Accessed December 22, 2025.
- Sleeman, Jennifer, Milton Halem, Tim Finin, and Mark Cane. 2017. «Modeling the Evolution of Climate Change Assessment Research Using Dynamic Topic Models and Cross-Domain Divergence Maps». <https://aaai.org/papers/15286-15286-modeling-the-evolution-of-climate-change-assessment-research-using-dynamic-topic-models-and-cross-domain-divergence-maps/>.
- Strittmatter, Anton. 1990. «Lehrerfortbildung in der Schweiz – Strukturen und Perspektiven». *Beiträge zur Lehrerbildung* 8 (1), S. 58–64. *Beiträge zur Lehrerbildung* 8. <https://doi.org/10.25656/01:13184>.
- Van Dijk, Jan. 2020. *The Digital Divide*. Cambridge: Polity.