Article

An NLP Approach for Extracting Practical Knowledge from a CMS-Based Community of Practice in E-Learning

Javier Leung [ID]

School of Information Science and Learning Technologies, University of Missouri, Columbia, MO 65201, USA; leungj@missouri.edu

Abstract: This study aimed to identify the tacit or practical knowledge of an online community of practice (CoP) based on a content management system (CMS) technology. The E-Learning Industry site is one of the most prominent news outlets that provides instructional design and technology (IDT) practitioners with insights into the field. Natural language processing (NLP) techniques were implemented to extract practical knowledge of publicly available and not password-protected text sources in seven news categories. First, the findings suggest emphasizing the production of online articles related to the production of e-learning materials in technology-enabled environments. Second, the results indicate the alternative uses of learning management systems to manage different aspects of the production of e-learning materials. Third, the findings show that the CoP's main priority was to reference existing materials in the community and external resources. The results of this study have implications and provide recommendations for researchers, community leaders, and practitioners toward improving knowledge discovery mechanisms, increasing transparency and integrity in communities, and increasing practitioners' ability to self-assess existing practical knowledge against competencies in the field. The present study takes an inventory of the organizational knowledge capital and functions embedded in a CoP using a CMS platform as a delivery mechanism for creating and sharing knowledge.

Keywords: topic modeling; data mining; instructional design; online learning; knowledge management; communities of practice

1. Introduction

Becoming an instructional designer requires formal training from curricular activities and practical experiences from internships that are similarly borrowed from other design fields, such as engineering and architecture [1]. Much of the literature on becoming an instructional designer has emphasized training methods to improve formal training. Though the literature offers formal methods of instructional design training in the form of design studio approaches, case studies, competitions, internships, and expert demonstrations that address the complexity of design problems and solutions, the practical aspects of instructional design work appear in the form of informal learning in the workplace [2–4] and educational experiences in adult and continuing education [5].

Information Communication and Technology (ICT) tools play an important role in supporting formal and informal learning activities through interconnected technological ecosystems that enable individuals to learn intentionally or unintentionally through Massive Online Open Courses (MOOCs), communities of practice (CoPs), and social media platforms [6]. Web 2.0 technologies allow individuals to create and distribute content [7]. Examples of Web 2.0 platforms are social networking sites (e.g., Facebook and LinkedIn), blog platforms (e.g., Blogger, Reddit, and WordPress), content and learning management systems (e.g., Drupal, Joomla, WordPress, Canvas, and Blackboard), and photo and video platforms (e.g., Instagram and YouTube).

Due to the rise of ICT tools and the open-source nature of content management systems (CMS), CoPs can deploy online communities based on Drupal, Joomla, or WordPress. Online CoPs can deliver informal learning opportunities and resources that address the needs and gaps of instructional designers [8]. As instructional designers expand their tacit or practical knowledge, informal learning experiences progressively allow practitioners to refine their skills over time. Informal learning is unplanned, unstructured, and incidental learning beyond formal settings and is not bound to a specific place and time [9,10]. Informal learning is also influenced by the presence or absence of the intentionality and consciousness of learning that can take place in the form of self-directed learning or implicit learning [11]. In self-directed learning, learners attempt learning activities that are conscious and intentional. In contrast, in implicit learning, learners are immersed in a context where they are not consciously trying to learn the subject.

The E-Learning Industry news outlet is an example of an online CoP that offers informal learning opportunities for learning the practical aspects of instructional design and technology (IDT). The E-Learning Industry site is based on a CMS that uses WordPress as a blog-publishing system to manage web content and users [12]. This online CoP has invited practitioners to publish online articles related to practical knowledge of the field since 23 February 2012 [13,14]. Their practical knowledge is organized into seven news categories that represent the general organization of online articles as follows: (1) Learning Management Systems, (2) E-Learning software, (3) E-Learning Trends, (4) Design and Development, (5) Instructional Design, (6) Best Practices, and (7) Free Resources. According to the website traffic report by Similarweb [15], the E-Learning Industry site attracts 389,312 unique monthly visitors, as of February 2022. The majority of the web traffic comes from the United States, India, the Philippines, the United Kingdom, and Canada. The organic keywords that generate free traffic to the site include *learning management system*, *advantages of online education*, *pros and cons of online learning*, and *forgetting curve*.

Studies in the knowledge management literature examine tacit knowledge extraction from explicit forms of knowledge in the workplace (e.g., online platforms, documents, and e-mail communication). These studies are explored through the SECI model, where knowledge is continuously created through socialization, externalization, combination, and internalization [16]. First, tacit knowledge is created through a socialization process, and its tacitness is difficult to codify into explicit knowledge. Second, tacit knowledge is externalized or articulated in symbolic language for sharing with other groups or individuals. Third, the combination step requires applying and reorganizing explicit knowledge. Fourth, when explicit knowledge is applied, individuals embody the knowledge as tacit through action and reflection. This present study aims to extract tacit or practical knowledge from explicit knowledge in text artifacts occurring at the externalization stage.

The contributions of this work include filling a gap in the IDT literature and creating a baseline for future studies that improve the mechanisms for sharing practical knowledge in alignment with professional competencies. First, the characteristics of practical knowledge among IDT CoPs in virtual environments are unknown in the IDT literature. While present studies examine instructional designers' professional development needs and roles in academic and corporate settings, exploring sources of practical knowledge in a virtual environment is required to understand the current knowledge structures and gaps in instructional design knowledge. Second, this study establishes a foundation for future studies that supports the development of intelligence and recommendation systems that allow practitioners to make better use of online resources for skill development and to detect misinformation about learning. The study explores the following research questions:

- RQ 1: What are the text characteristics, most frequent words, and word sequences used in the online community?
- RQ 2: What are the characteristics of sentiment, named entities, and relationships among entities in the online community?
- RQ 3: What are the latent topic structures in the online community?

The rest of the article is organized as follows. Section 2 provides a literature review and related studies. Section 3 describes the research methodology, including a thorough description of the natural language processing (NLP) tasks performed. Section 4 describes the results of NLP tasks organized by the research question. Section 5 contains a discussion of results, implications for research and practice, limitations, and recommendations for improving the CoP of interest. Finally, Section 6 concludes the article and provides the future direction of this work.

2. Background

Four concepts are essential to consider in the context of the study. These concepts include communities of practice, professional organizations in IDT, the characteristics and extraction of tacit or practical knowledge from unstructured or textual data, and related studies of online news sources using NLP.

2.1. Communities of Practice in Instructional Design

Wenger [17] coined the term *communities of practice* (CoP) to examine the learning among practitioners in a social environment. Wenger and Synder [18] described CoPs as groups of people informally gathered to share expertise in a specific domain or field as they interact regularly. Yanchar and Hawkley [1] argued that instructional design practitioners are willing to engage in informal learning efforts that address rapidly changing work situations. Online CoPs provide professional connections and supporting mechanisms with geographically dispersed members through ICT tools [19].

Schwier et al. [20] argued that instructional design CoPs are born of the convenience that allows informal engagement to solve specific project challenges or issues. The authors also investigated the features of instructional design CoPs in terms of history and culture, mutuality, plurality, and tacit knowledge. They found that shared history and culture are not prominent features in instructional design CoPs. In contrast, passive participation as a spectator was a critical element aligned with practitioners' agendas and community values. In terms of mutuality, community members developed their protocols for contribution and interaction with others. At the same time, community participation was based on the plurality of intermediate relationships with other members (i.e., experts in the field), which provided a wide range of considerations and solutions to learning problems.

Furthermore, Schwier et al. [21] investigated the types of agency or sense of responsibility that instructional designers hold in instructional design communities, their profession, and their respective work contexts. Interpersonal agency refers to one's capacity to exert control or influence the processes and outcomes of instructional design projects. Professional agency refers to the feeling of responsibility to the profession and community by acting in a professionally competent manner. Institutional agency refers to the sense of responsibility to advance the organization's agenda, which instructional designers represent. Societal agency is characterized by the sense of a contribution to society through instructional design work. Instructional design communities are also knowledge repositories where members can draw upon practical knowledge as members collectively transform tacit knowledge into explicit forms in informal and serendipitous ways. Examples of practical knowledge include unique or creative solutions to dealing with demanding clients, job aids or templates for applying criteria to projects, and expert advice to solve complex problems. Interestingly, the authors of [21] pointed out that healthy communities collaboratively rely on designing solutions to complex problems.

2.2. Instructional Design Competencies

Due to the absence of a recognized accrediting body that identifies the required competencies for IDT professionals, professional organizations have developed the competencies that define professionals' knowledge, skills, and abilities. Professional organizations use competencies to encapsulate professional benchmarks, responsibilities, and capabilities in different roles (e.g., training manager, evaluator, instructional designer, or instructional tech-

nologist). These competencies come from the American Talent Development (ATD) [22], the International Board of Standards for Training, Performance, and Instruction (IBSTPI) [23], the Association for Educational Communications and Technology (AECT) [24], and the International Society for Technology in Education (ISTE) [25].

2.3. Tacit Knowledge Characteristics and Extraction

Polanyi [26] initially introduced tacit knowledge with the assertion that "we know more than we can tell" regarding individuals' "know-how", "working knowledge", "expertise", or a set of abilities to perform a job that is difficult to articulate or transfer to others explicitly. Wagner and Sternberg [27] defined tacit knowledge as work-related practical knowledge learned informally through experience on the job, concerned with knowing how instead of knowing what.

McAdam et al. [28] stated that tacit knowledge has technical and cognitive dimensions that contain mental models, values, beliefs, and perceptions. Tacit cognitive knowledge incorporates implicit mental models and perceptions that allow individuals to understand their surroundings and tasks. Tacit technical knowledge is workers' knowledge and abilities to perform functions that are not easily articulated. Viale and Pozzali [29] argued that different forms of tacit knowledge could be acquired and transmitted in the form of competencies, background knowledge, and implicit cognitive rules, as defined below:

- Tacit knowledge as a competence refers to the skills and abilities acquired through apprenticeships and face-to-face interactions.
- Tacit background knowledge is the regulations, codes of conduct, and processes of acculturation to which individuals adhere, based on their context.
- Tacit knowledge acts as a mechanism for creating new knowledge and assessing the accuracy of information itself.

Steiger and Steiger [30] argued that tacit knowledge structures represent the implicit mental models of individuals. Mental models are tacit, where knowledge structures integrate the ideas, practices, assumptions, beliefs, relationships, facts, and misconceptions that individuals use to perceive and interact with others [31]. The authors [30] argued that tacit knowledge could be extracted from externalized knowledge through artificial neural networks and decision trees that perform the cognitive mapping of decision processes. Additionally, NLP algorithms are implemented to elicit, extract, and represent tacit knowledge from individuals and artifacts [32–38].

2.4. NLP Studies on Online News Sources

Online news outlets play a critical role in society, as a place where readers can learn about events, people, places, and trends. Additionally, online news outlets position themselves as credible information sources that provide readers with multiple perspectives on a given subject while producing information at a tremendous speed. However, studies of online news outlets using NLP have raised concerns about selectivity, misinformation, and transparency, which have had cognitive consequences for readers who encountered pieces of evidence, warrants, and claims to support persuasive written communication [39]. These studies have also identified the lack of mechanisms to improve the credibility, inclusivity, and fact-checking of online news articles. Ethical critiques of online news sources can identify the issues mentioned above through NLP techniques to detect the emerging contexts and classify online articles based on a given set of features. For instance, Srivastav and Singh [40] implemented a topic modeling approach to determine the contexts emerging from online news categories. Shang et al. [41] developed an application for auditing the production of published articles by detecting diversity, equity, and inclusion (DEI) indicators. Fung et al. [42] created an application for detecting users' sentiments and stances when reacting to news articles on social media. Singh and Singh [43] used a text similarity method to identify the issue of selectivity across various online news sources. Yu et al. [44] proposed a transformer-based machine learning technique for NLP to detect persuasion techniques in propagandistic news on social media. Gao et al. [45] performed supervised

and unsupervised NLP tasks on construction-related news outlets to classify and detect risk narratives. Jaidka et al. [46] implemented deep learning models to study persuasive communication in editing actions from Wikipedia Talk pages and predict editorial behavior and emotional change among contributors.

3. Materials and Methods

A total of 9033 online articles from April 2012 to September 2020 were scraped across seven news categories. The text sources from each news category were publicly available and required no password authentication to access the online articles. Each news category was identified based on the sitemap of the website to avoid the duplication of articles within and across categories. By obtaining all the links for each category, articles were scraped to include the title and body of the article, without author information. Table 1 shows the number of scraped articles for each news category.

Table 1. Number of online articles by news category.

News Category	Number of Articles
Learning Management System	927
E-Learning Software	400
E-Learning Trends	2934
Design and Development	2415
Instructional Design	1065
Best Practices	972
Free Resources	320
Total	9033

The average word and sentence lengths, word frequencies, and trigrams were generated as an exploratory step to understand the lengths of online articles, word frequencies, and probabilities of words appearing together. A stop words dictionary was not implemented in the average word and sentence lengths to count all words in the texts. In contrast, word frequencies and trigrams required a stop words dictionary to filter extraneous frequencies of common words, including articles, prepositions, pronouns, and conjunctions. After using a stop words dictionary, sentiment analysis, entity recognition, entity relationships, and topic modeling were implemented to extract sentiment polarity, pedagogical and educational technology entities and their relationships, and emerging themes from a news category. The remainder of this section describes the details of each NLP task performed in the study. Table 2 lists the Python packages used in the study.

Table 2. Summary of Python packages.

NLP Task	Python Package
Text characteristics	Lambda functions to calculate average word and sentence lengths
Visualization	Profile Report to visualize text characteristics
Sentiment analysis	TextBlob
Trigrams	NLTK
NER	spaCy
Topic modeling	Gensim for Latent Dirichlet Allocation and BERTopic (stsb-bert-large pre-trained model)

To address the first research question, the average word and sentence lengths of the online articles were generated using the lambda functions to explore text characteristics without filtering out stop words to account for all words. The generated features for each news category were visualized in the Profile Report package [47]. Additionally, word frequencies were visualized with the WordCloud package to identify prominent words in each news category [48]. Word frequencies were obtained after performing NLP tasks for cleaning, normalizing, and parsing using the Natural Language Toolkit (NLTK) by

performing lower casing, tokenization, stop word removal, lemmatization, stemming, and tagging parts-of-speech (POS). Though there is no consensus on using a standard stop word dictionary, the removal of stop words from textual data is a typical pre-processing step to remove noise or low-level information and reduce training time and dimensionality from uninformative words [49,50]. This study implemented the stop word English dictionary, Wordnet Lemmatizer, SnowBall Stemmer, and POS tagger libraries in NLTK. The NLTK n-gram language model package was implemented to create the probabilities of contiguous words in trigrams [51]. The most frequent trigrams were reported to illustrate word sequences in order to explore the context of the words.

To address the second research question, this study employed sentiment analysis, entity recognition, and entity relationships approaches. The TextBlob package was implemented for sentiment analysis to identify positive, neutral, and negative attitudes in the texts [52]. Online articles were classified as positive (1), neutral (0), and negative (-1). The spaCy package was implemented for the named entity recognition (NER) tasks to extract the names of people, places, organizations, and geographic locations [53]. Once entities were extracted, subject–object relationships emerged as entity pairs, allowing an understanding of how entities were referenced. These entity pairs consisted of the source and target entities linked by edge entities that defined the relationships among the entities.

In the third research question, the Latent Dirichlet Allocation (LDA) and BERTopic packages were used for topic modeling to discover latent topic patterns in each news category. In the first two rounds of topic modeling using LDA, the LDA topic modeling algorithm in the Gensim library generated word representations and probabilities using the bag-of-words (BoW) and Term Frequency–Inverse Document Frequency (TF-IDF) to predict emerging topic patterns in online articles from each news category [54]. The third model used sentence transformers with the BERTopic library using a class-based TF-IDF (c-TF-IDF) [55,56].

The LDA algorithm required a specific parameter for determining the exact number of topics that the algorithm used to achieve distinct and coherent topics. The ideal number of topics (n_topics) was achieved by running the LDA several times with multiple topic parameters from 2 to 20 until the elbow method achieved the highest coherence or C_v value. LDA also required parameters for the Dirichlet hyperparameter *alpha* for document–topic density and Dirichlet hyperparameter *beta* for word–topic density. The *alpha* and *beta* parameters were set to 'auto,' allowing the LDA algorithm to estimate the document–topic and word–topic densities automatically. With the BoW model, the TF-IDF model was generated to measure the importance of words against the whole corpus in the category. TF-IDF generated features or classes based on the term frequency of words and their weights in a document compared with their frequencies across all documents within a category.

In the third model, the BERTopic used sentence transformers and c-TF-IDF to calculate words' left and right contexts, generating clusters for topic interpretation. In c-TF-IDF, text sources are treated as a single class. Then, the frequency of each word was extracted and divided by the total number of words and documents across all classes. With BERTopic, the pre-trained sentence transformer model (stsb-bert-large) was implemented to identify semantic textual similarity by reducing dimensionality with the Uniform Manifold Approximation and Projection for Dimension Reduction technique (UMAP) and clustering sentence embeddings with the HDBSCAN algorithm [55].

Chang et al. [57] argued that there is no gold standard for evaluating topic models. Semantically coherent topics were examined through human judgment and quantitative approaches. Chang et al. [57] proposed evaluating topic model outputs using two methods, including topic and word intrusion methods. Regarding topic intrusion, discovered topics were evaluated based on whether the topic model's decomposition of the text sources agreed with human judgment based on domain expertise. In addition, a topic model was examined in terms of word intrusion by observing the words inserted in a topic model that did not provide semantic coherence or coherent meaning.

The semantic coherence values of topic models were also assessed quantitatively by obtaining semantic coherence measures, or *C_v* values. Semantic coherence measures describe how often topic words appear together in the corpus [58]. Table 3 summarizes the *C_v* values and the parameters for the ideal number of topics that resulted in semantically coherent topic models for each news category.

Table 3. *C_v* values and number of topics parameter for the LDA algorithm.

News Category	*C_v*	N_Topics Parameter
1. Learning Management System	0.488	7
2. E-Learning Software	0.437	4
3. E-Learning Trends	0.465	3
4. Design and Development	0.377	9
5. Instructional Design	0.438	5
6. Best Practices	0.422	3
7. Free Resources	0.356	10

Ethical Considerations

Even though web scraping is still a relatively new and emerging practice, Krotow and Silva [59] argued that ethical issues are associated with the automatic extraction of information. According to the authors, web scraping brings forth five ethical considerations: individual privacy and the rights of research participants, discrimination and bias, organization privacy, diminishing organizational value, and impacts on decision making. Even though web scraping involves ethical hurdles for academic researchers, and the Terms of Service (TOS) explicitly prohibit web scraping and the crawling of their platforms, Mancosu and Vegetti [60] noted that scraping public information from online platforms may be safe for researchers because research on social media serves the public interest. Additionally, Catanese et al. [61] argued that TOS is designed to disrupt the status quo by enforcing behavioral and technical limitations to web scraping.
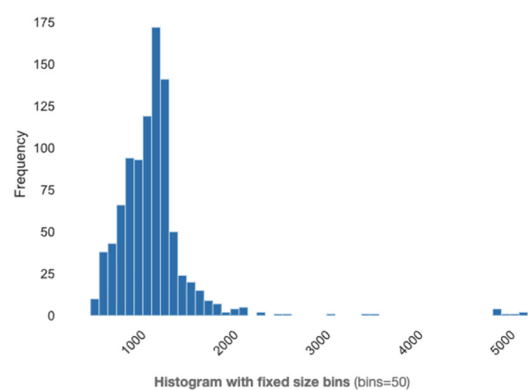
While technology plays a critical role in sustaining knowledge creation and sharing, technology can have negative consequences when comparing several online CoPs because of the lack of anonymity and privacy, which leads to the unintended identification of users (e.g., searching for articles or posts on their platforms) who choose to participate in communities. For this particular reason, any identifiable information (i.e., links to articles and authorship) was deleted to ensure the anonymity and privacy of authors from the e-learning news outlet. Text sources are not publicly available to prevent plagiarism and to protect the community's organizational knowledge [62].
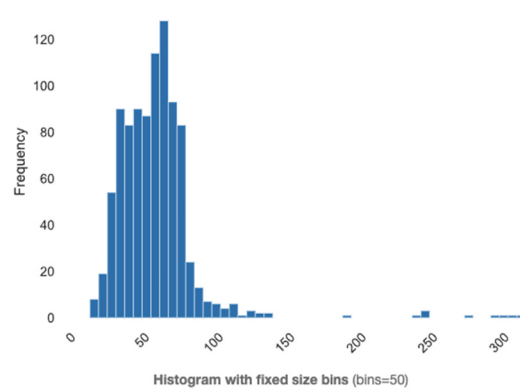
4. Results

4.1. RQ1: What Are the Text Characteristics, Most Frequent Words, and Word Sequences Used in the Online Community?
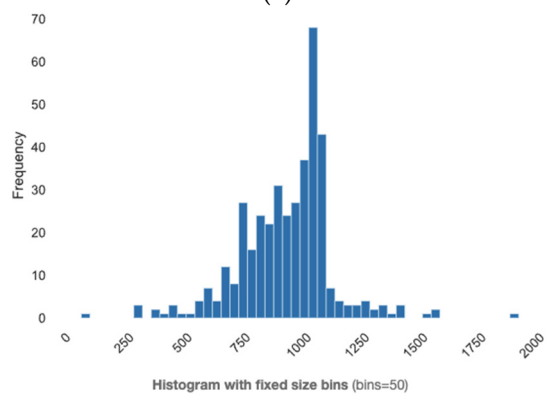
4.1.1. Text Characteristics

The below text characteristics accounted for all words in online articles without filtering stop words. In the Learning Management System category, the average word count was 1214.37 words, and the average sentence count was 58.75 sentences. The E-Learning Software category contained an average word count of 946.00 words and an average sentence count of 51.50 sentences. In the E-Learning Trends category, the average word count was 690.35 words, and the average sentence count was 37.51 sentences. The Design and Development category contained an average word count of 661.33 words and an average sentence count of 35.88 sentences. The Instructional Design category had an average word count of 681.25 words and an average sentence count of 36.76 sentences. In the Best Practices category, the average word count was 670.81 words, and the average sentence count was 36.68 sentences. In the Free Resources category, the average word count was 523.19 words, and the average sentence count was 28.25 sentences. The word and sentence length distributions are reported in Figure 1.
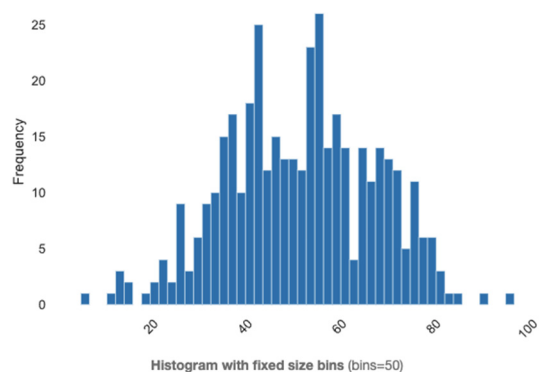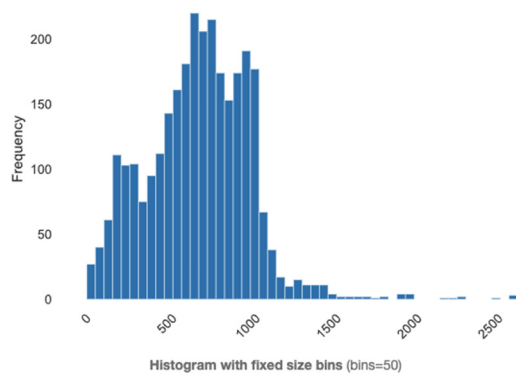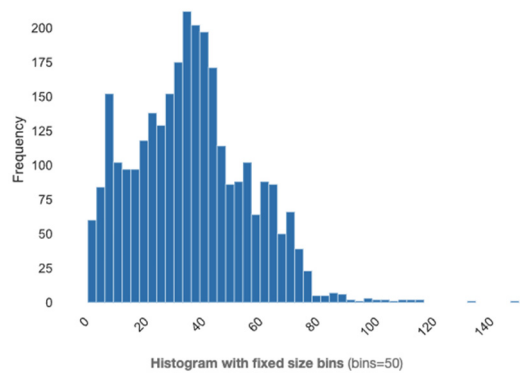
Figure 1. *Cont.*

(**g**)

(**h**)

(**i**)

(**j**)

(**k**)

(**l**)

(**m**)

(**n**)

Figure 1. Distributions of the word and sentence lengths in each news category: (**a**) Learning Management System avg. word length; (**b**) Learning Management Systems avg. sentence length; (**c**) E-Learning Software avg. word length; (**d**) E-Learning Software avg. sentence length; (**e**) E-Learning Trends avg. word length; (**f**) E-Learning Trends avg. sentence length; (**g**) Design and Development avg. word length; (**h**) Design and Development avg. sentence length; (**i**) Instructional Design avg. word length; (**j**) Instructional Design avg. sentence length; (**k**) Best Practices avg. word length; (**l**) Best Practices avg. sentence length; (**m**) Free Resources avg. word length; (**n**) Free Resources avg. sentence length.

4.1.2. Word Frequencies

After performing text processing and using a stop words dictionary, the most frequent words from the online articles emerged as unique tokens or words that were the most representative of the category. Based on the size of the dictionary after removing stop words, the following list ranks the number of unique words found in each category in descending order:

1. Category 3: E-Learning Trends (36,321 words);
2. Category 4: Design and Development (30,808 words);
3. Category 5: Instructional Design (19,615 words);
4. Category 6: Best Practices (19,102 words);
5. Category 1: Learning Management System (17,580 words);
6. Category 2: E-Learning Software (11,325 words);
7. Category 7: Free Resources (9481 words).

In first place, the E-Learning Trends category had the largest dictionary, and the three most frequent words were *learner* (7360), *need* (5958), and *use* (5122). In second place, the three most frequent words in the Design and Development category were *learner* (6269), *need* (5355), and *elearning course* (4976). In third place, the three most frequent words in the Instructional Design category were *learner* (2858), *learning* (2310), and *need* (2115). In fourth place, the three most frequent words in the Best Practices category were *learner* (2642), *student* (2588), and *course* (2107). In fifth place, the three most frequent words in the Learning Management System category were *need* (4799), *lms* (4699), and *will* (3356). In sixth place, the three most frequent words in the E-Learning Software category were *need* (1514), *online training* (1468), and *lms* (1465). In seventh place, the three most frequent words in the Free Resources category were *elearning* (678), *learning* (648), and *will* (590). Figure 2 shows the word cloud visualizations that represent word frequencies.

(a)   (b)

Figure 2. *Cont.*

Figure 2. Most frequent words in each news category: (**a**) Learning Management System; (**b**) E-Learning Software (**c**) E-Learning Trends; (**d**) Design and Development; (**e**) Instructional Design; (**f**) Best Practices; (**g**) Free Resources.
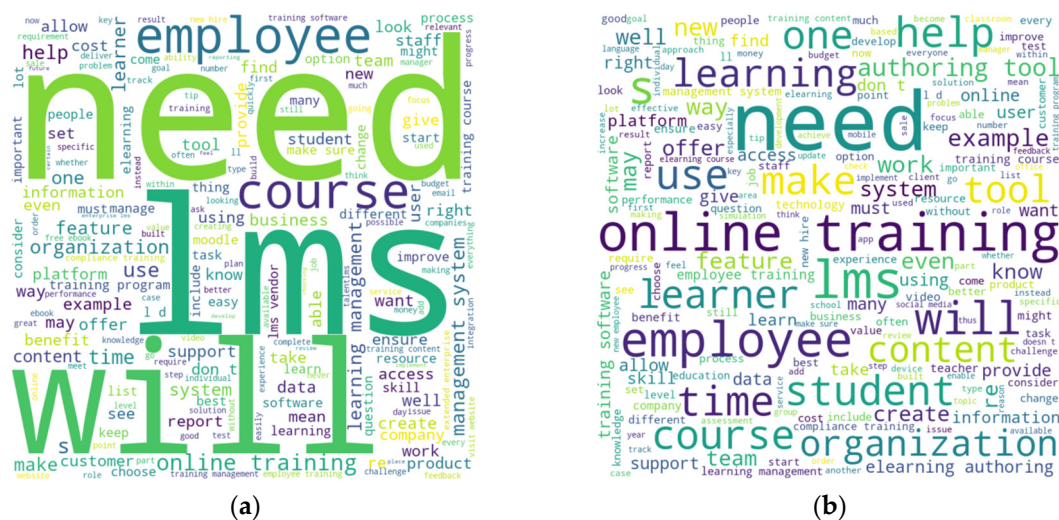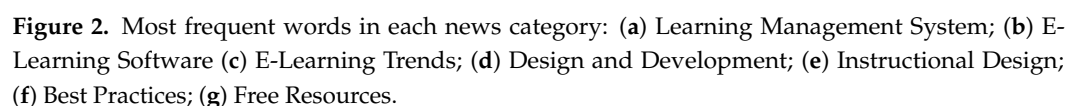
4.1.3. N-Grams

In the Learning Management System category, the most frequent trigrams were related to using learning management systems for tracking and managing time, running a small business, and being easy to implement and maintain. In the E-Learning Software category, the most frequent trigrams were related to rapid prototyping, the implementation of platforms in corporate settings, and custom e-learning development materials. In the E-Learning Trends category, the trigrams described instructional design theories and models, responsive LMS solutions, and support for mobile learning. In the fourth category, Design and Development, the trigrams suggested rapid e-learning authoring tools and extending LMS functionality with JavaScript and HTML code snippets. In the Instructional Design category, the most frequent trigrams were associated with instructional design models, theories and history, the discussion of alternative instructional design models, and graduate certificates in instructional design. In the sixth category, Best Practices, the most frequently occurring trigrams were related to professional guides in instructional design practice. In the last category, Free Resources, the trigrams were related to e-learning and educational technology tool tutorials and free resources. Table 4 summarizes the trigrams and their frequencies in each news category.

Table 4. Top 10 trigram frequencies in each category.

News Category	Trigram	Frequency
Learning Management System	[*learning management system*]	1511
	[*online training course*]	407
	[*extended enterprise lms*]	313
	[*running a small business*]	293
	[*help free tool*]	291
	[*homebase help free*]	291
	[*manage team visit*]	291
	[*business never harder*]	291
	[*time manage team*]	291
	[*make work easier*]	291
E-Learning Software	[*elearning authoring tool*]	406
	[*learning management system*]	319
	[*online training course*]	283
	[*online training content*]	169
	[*online training software*]	163
	[*employee training software*]	163
	[*online training resource*]	104
	[*lms training company*]	94
	[*employee training participant*]	87
	[*value money lms*]	82
E-Learning Trends	[*learning management system*]	609
	[*online training course*]	319
	[*subject matter expert*]	255
	[*instructional design model*]	252
	[*mobile learning strategy*]	209
	[*online training resource*]	207
	[*mobile learning solution*]	192
	[*elearning authoring tool*]	189
	[*elearning course design*]	169
	[*online training program*]	147

Table 4. *Cont.*

News Category	Trigram	Frequency
Design and Development	[*learning management system*]	409
	[*online training course*]	381
	[*elearning authoring tool*]	336
	[*elearning course design*]	321
	[*subject matter expert*]	275
	[*online training resource*]	192
	[*elearning content development*]	181
	[*elearning content provider*]	169
	[*online training content*]	156
	[*online training program*]	119
Instructional Design	[*instructional design model*]	252
	[*subject matter expert*]	180
	[*online training course*]	132
	[*elearning course design*]	119
	[*design model theory*]	114
	[*learning management system*]	107
	[*learning battle card*]	85
	[*online learner able*]	71
	[*give online learner*]	69
	[*elearning authoring tool*]	61
Best Practices	[*learning management system*]	143
	[*online training resource*]	92
	[*elearning course design*]	90
	[*online training course*]	90
	[*elearning authoring tool*]	90
	[*subject matter expert*]	65
	[*online training content*]	55
	[*curated elearning content*]	50
	[*online training program*]	49
	[*elearning content curation*]	47
Free Resources	[*learning management system*]	49
	[*free video tutorial*]	36
	[*elearning course design*]	33
	[*elearning infographic template*]	29
	[*camstasia studio 8*]	28
	[*top elearning blog*]	27
	[*adobe captivate 7*]	22
	[*elearning authoring tool*]	21
	[*free moodle video*]	20
	[*mobile apps learning*]	18

4.2. RQ2: What Are the Characteristics of Sentiment, Named Entities, and Relationships among Entities in the Online Community?

4.2.1. Sentiment

The majority of online articles had a positive sentiment across all categories. However, a few articles had neutral and negative sentiments, in all news categories except the Learning Management System category. Table 5 summarizes the sentiment distributions for each news category.

Table 5. Sentiment distribution of articles.

News Category	Positive	Neutral	Negative	Total
Learning Management System	927	0	0	927
E-Learning Software	399	0	1	400
E-Learning Trends	2908	10	16	2934
Design and Development	2404	5	6	2415
Instructional Design	1061	0	4	1065
Best Practices	964	3	5	972
Free Resources	318	1	1	320
Total	8981	19	33	9033

4.2.2. Recognized Entities

After performing entity recognition with spaCy, the recognized entities emerged as pedagogical and educational technology entities. Figure 3 shows the distributions of the most frequent entities by news category. Based on the entities recognized, the following list ranks the news category in descending order:

1. Category 3: E-Learning Trends (58,606 entities);
2. Category 4: Design and Development (50,326 entities);
3. Category 1: Learning Management System (35,747 entities);
4. Category 5: Instructional Design (24,097 entities);
5. Category 6: Best Practices (18,944 entities);
6. Category 2: E-Learning Software (12,562 entities);
7. Category 7: Free Resources (6939 entities).

In first place, the most frequently recognized entities in the E-Learning Trends category were *eLearning* (9128), *LMS* (1809), *Instructional Design* (823), *L&D* (679), *Instructional Designers* (432), *mLearning* (256), *ADDIE* (206), *YouTube* (183), *Learning Management System* (173), *an Instructional Designer* (134), *Learning Management Systems* (119), and *Instructional Designer* (93). Regarding concepts, frameworks, theories, technologies, and practitioners, less frequent entities were found as follows: *xAPI* (40), *Bloom* (36), *Gardner* (24), *Pavlov* (29), *Hermann Ebbinghaus* (20), *Kirkpatrick* (19), *Connie Malamed* (16), *Michael Allen* (15), *Reusable Learning Objects* (14), *Knowles* (13), *Gagne* (11), and *Kolb* (10).

In second place, the most frequently recognized entities in the Design and Development category were *eLearning* (15,674), *LMS* (719), *L&D* (581), *PowerPoint* (258), *Instructional Design* (241), *eBook* (237), *Instructional Designers* (200), *Elucidat* (123), *YouTube* (110), *SCORM* (109), *HTML5* (98), and *Learning Management System* (89). Less frequent entities were identified among concepts, frameworks, technologies, and theories as follows: *Flash* (72), *ADDIE* (61), *Project Management* (55), *Camtasia* (51), *Bloom* (32), *Reusable Learning Objects* (28), and *Instructional Design* (20).

In third place, the most frequently recognized entities in the Learning Management System category were *LMS* (10,225), *eLearning* (1696), *Moodle* (466), *L&D* (462), *Learning Management System* (342), *Learning Management Systems* (246), *eCommerce* (123), *SME* (115), *a Learning Management System* (113), *CMS* (107), *eBook* (105), and *the Learning Management System* (91). Less frequent entities were Section 508 (2) and *WCAG 2.0* (2), referring to concepts, frameworks, and theories related to instructional design.

In fourth place, the most frequently recognized entities in the Instructional Design category were *eLearning* (4494), *Instructional Design* (760), *Instructional Designers* (337), *ADDIE* (200), *LMS* (165), *L&D* (161), *an Instructional Designer* (123), *eBook* (85), *the Instructional Designer* (83), *Instructional Designer* (82), *SME* (64), and *Design Thinking* (39). Less frequent entities in the Instructional Design category were related to models, concepts, and theories, including *SAM* (34), *Bloom* (32), *Project Management* (30), *Pavlov* (28), *Merrill* (26), *Malcolm Knowles* (17), *Sweller* (15), *Dick & Carey* (15), *Reusable Learning Objects* (14), *Skinner* (12), *Cognitive Apprenticeship Model* (12), *Howard Gardner* (11), *Herman Ebbinghaus* (11), *Vygotsky* (11), *Rapid Prototyping* (9), *Elaboration Theory* (4), *The Agile Manifesto* (3), *Cognitive Load Theory* (3), *The Spiral Model* (3), *T-Shaped Learning Design Interest Approach* (2), *Nine Events of*
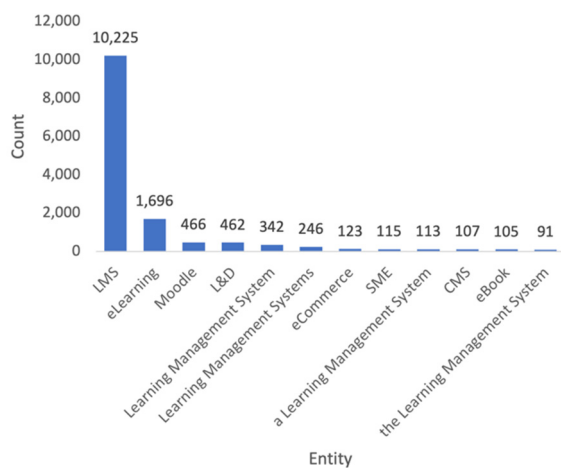
Instruction (2), *ARCS model of Motivation* (2), *Collaborative Learning Approach* (2), *Discovery Learning Model* (2), *AGES Model* (2), and *Basic Action Workflow* (2).
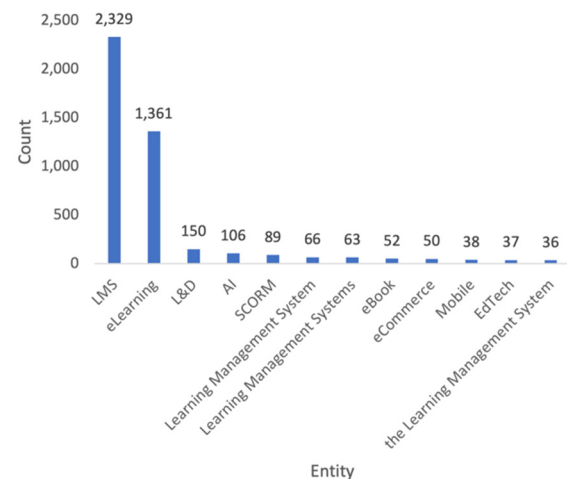
In fifth place, the most frequently recognized entities in the Best Practices category were *eLearning* (4717), *LMS* (409), *PowerPoint* (89), *L&D* (80), *Instructional Design* (67), *YouTube* (60), *Instructional Designers* (56), *eBook* (48), *Learning Management System* (40), *Elucidat* (37), *eLearners* (34), and *PDF* (32). Regarding models, concepts, and theories, less frequent entities were identified, as follows: *Principles of Effective Online Pedagogy* (2), *The Importance Of Meaningful Online Feedback* (2), *Active Learning* (2), *Bernard* (2), *Instructional Design Model* (1), Section 508 (1), and *American Disabilities Act* (1).

In sixth place, the most frequently recognized entities in the E-Learning Software category were *LMS* (2329), *eLearning* (1361), *L&D* (150), *AI* (106), *SCORM* (89), *Learning Management System* (66), *Learning Management Systems* (63), *eBook* (52), *eCommerce* (50), *Mobile* (38), *EdTech* (37), and *the Learning Management System* (36). Less frequent entities in this category involved *JIT* (just in time, 33), *LXP* (learning experience platform, 21), *eBooks* (17), *ADDIE Model* (1), Section 508 (3), *Universal Design for Learning* (1), and *Learning Methods* (1). The ebooks were related to corporate training, new employee onboarding, and the branding of online courses.

In seventh place, the most frequently recognized entities in the Free Resources category were *eLearning* (1414), *LMS* (119), *eBook* (63), *L&D* (59), *Adobe Captivate* (26), *Instructional Design* (25), *PowerPoint* (24), *eLearning Infographics* (23), *eLearning Industry* (22), *Mobile* (21), *Camtasia* (21), and *Learning Management Systems* (19). In this particular category, less frequent entities were *Docebo* (16), *eBooks* (15), *Adobe Captivate* (14), *Blackboard* (8), *Camtasia Studio* (7), *Vidopop* (7), *Kallidus* (6), *Snagit* (6), *Docebo* (5), and *Nine Events of Instruction* (1). In this category, the ebooks offered different topics related to managing learning objects, vendors for learning technologies, free infographic tools, and how to become an instructional designer.

(**a**)

(**b**)

Figure 3. *Cont.*

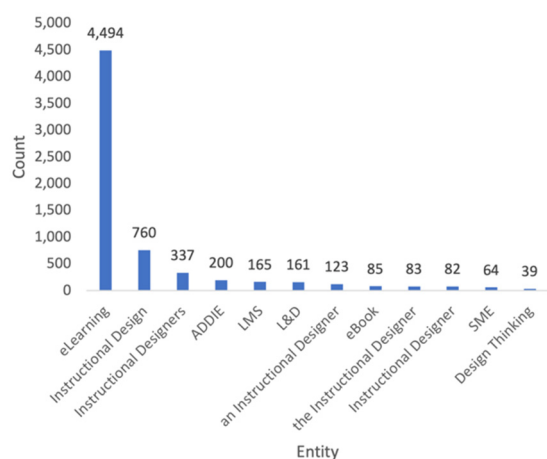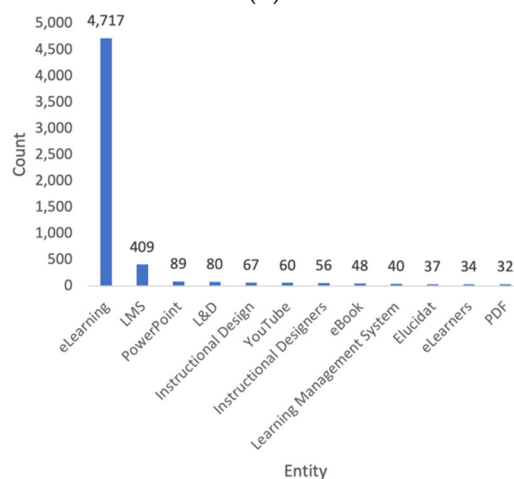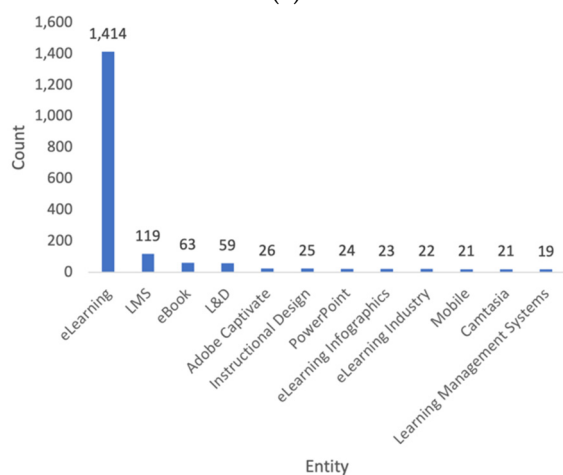**Figure 3.** Most recognized entities in each news category: (**a**) Learning Management System; (**b**) E-Learning Software (**c**) E-Learning Trends; (**d**) Design and Development; (**e**) Instructional Design; (**f**) Best Practices; (**g**) Free Resources.

4.2.3. Entity Relationships

Prominent relationships between entities were extracted when community members used specific words to describe pedagogical and educational technology entities. The words used to define the relationships between entities identified how the community described the context of pedagogical and educational technology elements in the online articles. Table 6 summarizes the 10 most frequent entity relationships for each news category. Based on the entity relationships recognized, the following list ranks the news categories in descending order:

1. Category 3: E-Learning Trends (1326 entity relationships);
2. Category 4: Design and Development (1174 entity relationships);
3. Category 6: Best Practices (626 entity relationships);
4. Category 5: Instructional Design (611 entity relationships);
5. Category 2: E-Learning Software (300 entity relationships);
6. Category 7: Free Resources (236 entity relationships);
7. Category 1: Learning Management System (213 entity relationships).

In first place, the E-Learning Trends category had 1326 entity relationships that demonstrated the ease of use of learning management systems (e.g., Talent LMS, Administrate) for deploying online courses aimed at users with zero experience, using simple interfaces. In second place, the Design and Development category identified 1174 entity relationships that emphasized strategies and tips for developing effective e-learning (e.g., learner engagement and motivation strategies), the formative assessment of online courses, and free media resources.

In third place, the Best Practices category contained 626 entity relationships that suggested project management strategies for e-learning development, instructional strategies (e.g., gamification and avatars), effective feedback to learners, and the creation of high-quality images. In fourth place, the Instructional Design category had 611 entity relationships that suggested creating memorable e-learning experiences and incorporating learning theories in e-learning development (e.g., cognitive load multimedia learning, adult learning, active learning, and Ebbinghaus' forgetting curve).

In fifth place, the E-Learning Software category identified 300 entity relationships that suggested several aspects of learning management systems for training, reporting, aligning e-learning materials with business goals, and licensing software options. In sixth place, the Free Resources category contained 236 entity relationships that offered free Web 2.0 resources, free stock photo libraries, and webinar and conference resources. In seventh place, the Learning Management System category had 213 entity relationships, including reviews of learning management systems (e.g., Talent LMS, ShareKowledge, Administrate), administrative and learner features, and implementation costs for organizations.

Table 6. Top 10 most frequent entity relationships and frequencies by news category.

News Category	Entity	Frequency
Learning Management System	*website*	509
	read	67
	published	35
	want	14
	captivate prime	12
	means	8
	features	6
	allow	5
	help	5
	take free	5

Table 6. *Cont.*

News Category	Entity	Frequency
E-Learning Software	*studio*	8
	references	6
	costs	5
	use	4
	features	4
	take	4
	conclusion	4
	need	3
	help	3
	halfpoint	2
E-Learning Trends	*read*	77
	need	25
	used	24
	want	22
	check	21
	use	21
	help	21
	leave	20
	studio	20
	find	18
Design and Development	*read*	107
	use	33
	help	27
	find	23
	take	20
	make	19
	create	18
	need	15
	professional	15
	keep	15
Instructional Design	*read*	41
	leave	12
	used	11
	think	9
	check	9
	find	9
	images	8
	offer instructional	8
	know	7
	learning	7
Best Practices	*read*	35
	use	13
	images	12
	help	11
	make	10
	want	8
	studio	8
	learn	7
	let	7
	find	7

Table 6. *Cont.*

News Category	Entity	Frequency
Free Resources	*join free*	6
	use	5
	visit	5
	find	4
	help	4
	read	4
	missed free	4
	captivate	4
	see	4
	want	3

4.3. RQ3: What Are the Latent Topic Structures in the Online Community?

Across all news categories, the BoW topic models had better topic interpretation than the TF-IDF models, based on the subject matter and higher probabilities of topic distributions. The BoW and sentence transformer topic models are described below by news category.

4.3.1. Category 1: Learning Management System

In the topic models using the BoW and sentence transformers, the Learning Management System category showed seven emerging topics related to custom training development, platform implementation costs and experiences, and the various uses of platforms for employee onboarding, online compliance training, and user and content management. Table 7 summarizes the BoW and sentence transformer topic models for the category.

Table 7. Emerging topic patterns for the Learning Management System category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	Employee Onboarding	Employee Training
Topic 2	Custom Online Training	Online Compliance Training
Topic 3	LMS Requirements	LMS User and Content Management
Topic 4	Online Compliance Training	LMS Implementation
Topic 5	LMS Implementation	Employee Training
Topic 6	Employee Training	LMS Implementation
Topic 7	LMS User and Content Management	Online Employee Training Costs

4.3.2. Category 2: E-Learning Software

In the BoW topic model, four topics were related to the various uses of e-learning development for employee onboarding, language acquisition courses, online compliance training, and e-learning software reviews. In the sentence transformer topic model, eight topic models were associated with platform use in various settings (e.g., K-12 and corporate), organizational goals (e.g., compliance and education), mobile learning support, user and content management, and platform requirements. Table 8 summarizes the BoW and sentence transformer topic models for the category.

Table 8. Emerging topic patterns for the E-Learning Software category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	Language Courses	LMS Requirements
Topic 2	Employee Onboarding	Employee Onboarding
Topic 3	Online Compliance Training	Technology in Educational Settings
Topic 4	E-Learning Authoring Tools	LMS in Corporate Settings
Topic 5		Technology in Educational Settings
Topic 6		LMS User and Content Management
Topic 7		Mobile Learning
Topic 8		E-Learning Authoring Tools

4.3.3. Category 3: E-Learning Trends

In the BoW topic model, three topic models described gamification, e-learning development, and mobile learning. Interestingly, the sentence transformer topic model generated 27 emerging topics. Five topics displayed the highest probability, including the instructional design process, microlearning, new employee onboarding, social media for collaboration and networking, and learning theories. Table 9 summarizes the BoW and sentence transformer topic models for the category.

Table 9. Emerging topic patterns for the E-Learning Trends category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	Gamification	Instructional Design Process
Topic 2	E-Learning Development	Microlearning
Topic 3	Mobile Learning	Collaboration and Networking
Topic 4		Employee Onboarding
Topic 5		Learning Theories

4.3.4. Category 4: Design and Development

In the BoW topic model, nine emerging topics were related to the assessment and various processes of e-learning and employee training development, including mobile learning, e-learning templates, e-learning examples, engagement strategies, the translation of courses, and voiceover recording. The sentence transformer topic model showed six topics related to educational animation, learning objectives development, the translation of courses, learning theories, assessment, and voiceover recording. Table 10 summarizes the BoW and sentence transformer topic models for the category.

Table 10. Emerging topic patterns for the Design and Development category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	Mobile Learning	Educational Animation
Topic 2	E-Learning Development	Course Translation
Topic 3	E-Learning Templates	Assessment
Topic 4	Employee Training	Voiceover
Topic 5	Voiceover	Learning Theories
Topic 6	E-Learning Examples	Learning Objectives
Topic 7	Engaging E-Learning	
Topic 8	Course Translation	
Topic 9	Assessment	

4.3.5. Category 5: Instructional Design

In the BoW topic model, two out of five similar topics were related to e-learning development, whereas the remainder were related to estimating time for e-learning development, employee training, and learning theories. The sentence transformer topic model generated 11 topics. Only six contained the highest probabilities, including e-learning development, learning theories, adult learning, instructional video development, user interface design, and instructional design jobs. Table 11 summarizes the BoW and sentence transformer topic models for the category.

Table 11. Emerging topic patterns for the Instructional Design category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	Estimate Development Time	E-Learning Development
Topic 2	Employee Training	Learning Theories
Topic 3	E-Learning Development	Adult Learning
Topic 4	E-Learning Development	Video Development
Topic 5	Learning Theories	User Interface Design
Topic 6		Instructional Design Jobs

4.3.6. Category 6: Best Practices

The BoW topic model described three emerging topics, with two similar topics relating to e-learning development, and one topic related to the translation of courses. In the sentence transformer topic model, four topics contained online learning, the development of language courses, student feedback, and instructional video development. Table 12 summarizes the BoW and sentence transformer topic models for the category.

Table 12. Emerging topic patterns for the Best Practices category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	E-Learning Development	Online Learning
Topic 2	Course Translation	Language Courses
Topic 3	E-Learning Development	Student Feedback
Topic 4		Video Development

4.3.7. Category 7: Free Resources

The BoW topic model showed 10 emerging topics related to different aspects of e-learning development (e.g., image and video editing, storyboarding), online resources (e.g., online webinars, image resources, and professional development opportunities), and e-learning authoring tips (e.g., Adobe Captivate and Camtasia). In the sentence transformer topic model, four emerging topics described online training resources, tips for e-learning authoring tools, e-learning development resources, and mobile apps. Table 13 summarizes the BoW and sentence transformer topic models for the category.

Table 13. Emerging topic patterns for the Free Resources category.

Topic	Bag-of-Words	Sentence Transformers
Topic 1	Multimedia Resources	Training Resources
Topic 2	Video Development	Adobe and Camtasia
Topic 3	Adobe Captivate	Infographic Resources
Topic 4	Employee Training	Apps
Topic 5	Storyboarding	
Topic 6	Professional Development	
Topic 7	Infographics	
Topic 8	Webinars	
Topic 9	Multimedia Resources	
Topic 10	Multimedia Resources	

5. Discussion

CMSs facilitate the creation and dissemination of knowledge. As a blog-publishing platform, the e-learning news outlet acts as a crowdsourcing mechanism for producing and managing online articles. The online articles from this CoP offer practitioners informal learning opportunities with easy-to-read online articles. Online articles in the Free Resources category had the shortest average word count of 523.19. In contrast, online articles in the Learning Management System and E-Learning Software categories had the highest average word counts of 1214.37 and 946, respectively. Online articles in the five remaining news categories had an average word count between 523.19 and 690.35. Although online articles in the Learning Management System and E-Learning Software categories were longer, this community of practice may require the writing of persuasive online articles to support educational technology vendors and recommendations. As Chambliss and Garner [39] stated, crafting convincing messaging requires three essential pieces, including the evidence, claim, and warrant, in order to sustain persuasive written communication in support of the author's arguments and lead to changed beliefs.

Though practitioners can submit online articles, the site did not describe the editing and review processes behind selecting credible sources of information. Interestingly, the

sentiment distribution of online articles indicated that most articles across the categories were positive. Only a few online articles had neutral and negative sentiments across six news categories, except the Learning Management System category. Positive online articles may be written persuasively to convince instructional designers and e-learning developers to adopt certain pedagogical practices and educational technology tools. Chambliss and Garner [39] argued that readers could change their beliefs while reading a persuasive text. Still, readers tended to revert to old ideas consistent with past experiences and background knowledge.

Regarding e-learning, this IDT community may have carefully crafted online articles in a positive tone that attract newly minted and accidental practitioners who are more willing to accept pedagogical and technical advice than experienced practitioners. Acevedo and Roque [63] argued that the instructional design field is at risk of deprofessionalization, resulting in non-experts becoming practitioners who prioritize online courseware production over learning theory, instructional design models, and pedagogy. Non-experts are practitioners who landed in instructional design and e-learning development roles, with the job of training others in the organization, without formal training [64]. While the E-Learning Industry site emphasizes the production of e-learning materials in their online articles, this CoP needs to associate technology-related articles with the instructional design news category. This way, practitioners of different professional backgrounds have the necessary pedagogical foundations to support technology-enabled learning environments and combat misconceptions about learning (e.g., learning styles).

5.1. Priorities of the Online CoP

The knowledge-production capability of this IDT community showcases the sense of responsibility (i.e., professional agency) to practitioners who access the site's resources by offering recommendations, evaluations, and reports that support the pedagogical and educational technology aspects of instructional design practice. The word frequencies emphasized learning management systems as a critical medium for managing and delivering e-learning courses to learners, students, or employees for online onboarding, skill development, and compliance reasons. The findings also suggested that the words *need*, *lms*, and *learner* were present across the seven categories with various degrees of frequency.

By observing the trigrams generated from each category, the findings point out the priorities of this online community. The use of learning management systems and online training courses was present across the seven news categories. Furthermore, the trigrams in the Learning Management System category showed the alternative uses of platforms for running e-learning shops as small businesses and managing e-learning development time. The trigrams in the E-Learning Software category emphasized the rapid development of online experiences, the integration of courses in learning management systems, and implementation costs. The trigrams in the E-Learning Trends category were characterized by authoring online training courses or training for mobile devices. The trigrams in the Instructional Design category were related to instructional design models, working with subject matter experts, e-learning course design processes, and gamified tools (e.g., learning battle cards) for learning instructional design and e-learning. The trigrams in the Best Practices category had a mix of overlapping trigrams from the first five news categories. The trigrams in the Free Resources category involved tutorials and templates to support practitioners with e-learning authoring tools (e.g., Camtasia, Captivate, and Moodle) and mobile apps.

Furthermore, the most frequently recognized entities were *eLearning*, *LMS*, and *L&D* (Learning and Development) across the seven categories. The predominant entity relationships were found in educational technology tools and e-learning platforms. The word *read* was the most common entity that tied educational technology and pedagogical entities across all categories, except the E-Learning Software and Free Resources categories. The entity relationships for *read* suggested that the authors of the articles tended to make refer-

ences to existing materials in the community or pointed to external resources (e.g., blog posts, research articles, and vendors).

5.2. E-Learning Materials Production as the Purpose of the Community

Based on the findings, the topic models suggested emphasizing the production of e-learning materials as the main purpose of the online community. In the Learning Management and E-Learning Software categories, most topic structures suggested the use of cases involved in training and the implementation of learning management systems and the authoring of online courses on platforms. In the E-Learning Trends category, topic models suggested a few instructional strategies, including gamification, social learning, and mobile learning. In the Design and Development category, the topic models emphasized e-learning development templates for online training, interactivity in e-learning (e.g., voice-over and animation), and the assessment of learners. In the Instructional Design category, most topic models were similar to those of the Design and Development category. A few topics in instructional design theories and user interface design were also present in the Instructional Design category. In the Best Practices category, the topic models were closely related to the Learning Management System and Design and Development categories. The Free Resources category had the highest number of topic models related to resources that support specific technical aspects of e-learning development.

5.3. Implications for Research and Practice

The results of the study have implications for researchers, practitioners, and leaders of online communities of practice. The findings of this study highlight the need to understand how professional competencies align with the community's practical knowledge based on the discovered topic models and pedagogical and educational technology entities. Further investigation of how practical knowledge in e-learning materials production is applied or transferred to practitioners' work contexts is required. Besides studying the learning transfer of informal learning opportunities, the quality of practical knowledge from the community needs to be further investigated to understand the inner workings and dynamics of the community's knowledge-production capabilities.

The findings suggest the need for practitioners to better understand how their self-conscious learning efforts integrate with their existing knowledge, skills, and abilities. Though this community is focused on the technical aspects of the production of e-learning materials, practitioners need to self-assess their personal knowledge-management capabilities to explore the required tacit cognitive knowledge (e.g., learning theories) required for integrating learning experiences in technology-enabled settings. The findings were mainly in line with those of North et al. [65]. The researchers showed that job postings in instructional design emphasized online training technology application and production. In contrast, capabilities in knowledge management, lifelong learning, and business insight received less attention.

As practitioners seek informal learning opportunities, leaders of online communities of practice can benefit from the findings by understanding the current capabilities for producing and sharing knowledge. Furthermore, leaders of these communities can assess their current organizational knowledge and competencies to provide additional informal learning opportunities in the learning sciences. Community leaders can also use the findings to develop mechanisms for improving the navigation and alignment of produced knowledge for instructional design competencies. In return, practitioners can evaluate their knowledge, skills, and abilities against the expectations and competencies of the profession.

5.4. Recommendations

Though the representation of practical knowledge was modeled using NLP, the inner workings of how the community selects and reviews online articles could not be observed. Four recommendations are necessary to sustain online communities and promote member

participation by increasing transparency, aligning with competencies, reusing knowledge, and establishing clear boundaries.

While practitioners can submit online articles to the online community, the first recommendation is to make submission and review requirements visible in the community by establishing a rubric to control contribution quality. The second recommendation involves providing members with mechanisms to align online articles with instructional design competencies. By providing such mechanisms, members are better positioned to self-assess knowledge, skills, and abilities. Due to the wealth of resources across several news categories, the third recommendation is to allow community members to browse related content across news categories. While the organizational scheme is useful for organizing online articles, members should be able to cross-check information across the categories to understand the different aspects of the profession. For example, when reviewing online articles related to the production of e-learning materials, members are presented with related content on how to support learner engagement and assessment. The fourth recommendation is to establish clear boundaries that reject self-serving online articles that deter productive participation, and that genuinely advance the field of instructional design. At the same time, community members can observe how the community behaves with integrity, especially new members participating on the periphery.

5.5. Limitations

The present study was not without limitations. While online articles contained videos and links to external resources, these online artifacts were not analyzed because they were hosted outside of the online community. Additionally, e-books and guides created by the community were not analyzed because these were referenced in the Free Resources category. Additionally, users' comments on articles were not extracted because these were not present across all online articles. A significant amount of tacit knowledge was contained in these external resources, but they were not analyzed due to time constraints and the additional processing time required to model additional data.

6. Conclusions

This study aimed to identify the tacit knowledge from an e-learning news outlet called elearningindustry.com. Practitioners were invited to write online articles under multiple categories. The E-Learning news outlet relies on a CMS where online articles are deployed under different categories and approved by community administrators. By examining the codified, or explicit knowledge, from online news articles, the study aimed to quantify the organizational knowledge capital of the community, the types of tacit knowledge, and the hidden topic structures present in each news category. The quantification of practical knowledge shows the knowledge-creation capabilities and priorities of this CoP. The findings suggest that most topics were related to tacit technical knowledge of the production of e-learning materials. Though tacit technical knowledge was prevalent, tacit cognitive knowledge was present to a lesser degree. The findings provide evidence of the types of practical knowledge that practitioners may use for their informal learning endeavors across several industries. The results offer topic-organization schemes to leaders of CoPs, in order to enhance practitioners' abilities to self-assess and organize their practical knowledge.

The future direction of this research involves an investigation of the emerging patterns of practical knowledge by the type of sentiment, in order to uncover the challenging aspects of instructional practice. Additionally, the findings of this study enable the future development of ontologies and taxonomies to classify types of practical knowledge (i.e., cognitive and technical) and make distinctions between pedagogical and educational technology entities, which can be applied to assess the practical knowledge present in other online CoPs.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The online news articles in this study are available from https://elearningindustry.com/ (accessed on 20 September 2020).

Conflicts of Interest: The author declares no conflict of interest.

References

1. Yanchar, S.C.; Hawkley, M.N. Instructional design and professional informal learning: Practices, tensions, and ironies. *J. Educ. Technol. Soc.* **2015**, *18*, 424–434.
2. Clinton, G.; Rieber, L.P. The Studio experience at the University of Georgia: An example of constructionist learning for adults. *Educ. Technol. Res. Dev.* **2010**, *58*, 755–780. [CrossRef]
3. Ertmer, P.A.; Stepich, D.A.; York, C.S.; Stickman, A.; Wu, X.; Zurek, S.; Goktas, Y. How instructional design experts use knowledge and experience to solve ill-structured problems. *Perform. Improv. Q.* **2008**, *21*, 17–42. [CrossRef]
4. Hardré, P.L.; Ge, X.; Thomas, M.K. An Investigation of Development Toward Instructional Design Expertise. *Perform. Improv. Q.* **2008**, *19*, 63–90. [CrossRef]
5. Lim, D.H.; You, J.; Kim, J.; Hwang, J. Instructional design for adult and continuing higher education: Theoretical and practical considerations. *Res. Anthol. Adult Educ. Dev. Lifelong Learn.* **2021**, 1018–1038. [CrossRef]
6. García-Peñalvo, F.J. Informal Learning Management Experiences. *Int. J. Hum. Cap. Inf. Technol. Prof.* **2014**, *5*, iv–ix.
7. Önday, Ö. Web 6.0: Journey from Web 1.0 to Web 6.0. *J. Media Manag.* **2019**, *1*, 1–6. [CrossRef]
8. Martinez, S.; Whiting, J. Designing Informal Learning Environments. In *Design for Learning: Principles, Processes, and Praxis*; McDonald, J.K., West, R.E., Eds.; EdTech Books: Online, 2021; Available online: https://edtechbooks.org/id/designing_informal (accessed on 6 March 2022).
9. Abramenka-Lachheb, V.; Lachheb, A.; Leung, J.; Sankaranarayanan, R.; Seo, G.Z. Instructional Designers' Use of Informal Learning: How Can We All Support Each Other in Times of Crisis? *J. Appl. Instr. Des.* **2021**, *10*. [CrossRef]
10. Conlon, T.J. A review of informal learning literature, theory and implications for practice in developing global professional competence. *J. Eur. Ind. Train.* **2004**, *28*, 283–295. [CrossRef]
11. Evans, J.R.; Karlsven, M.; Perry, S.B. Informal Learning. In *The Students' Guide to Learning Design and Research*; Kimmons, R., Ed.; EdTech Books: Online, 2018; Available online: https://edtechbooks.org/studentguide/informal_learning (accessed on 20 March 2022).
12. Detect which CMS a Site is Using-What CMS? What CMS Is This Site Using? Available online: https://whatcms.org/?s=elearningindustry.com (accessed on 18 March 2022).
13. elearning Industry Inc. About. Elearning Industry. Available online: https://elearningindustry.com/about-us (accessed on 18 March 2022).
14. ICANN. Registration Data Lookup Tool. Available online: http://maintenance.icann.org/lookup (accessed on 18 March 2022).
15. Similarweb. Available online: https://www.similarweb.com/ (accessed on 28 February 2022).
16. Nonaka, I.; Takeuchi, H. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*; Oxford University Press: Oxford, UK, 1995.
17. Wenger, E. Communities of practice: Learning as a social system. *Syst. Think.* **1998**, *9*, 2–3. [CrossRef]
18. Wenger, E.C.; Snyder, W.M. Communities of practice: The organizational frontier. *Harv. Bus. Rev.* **2000**, *78*, 139–146.
19. Gray, B. Informal learning in an online community of practice. *Int. J. E-Learn. Distance Educ./La Revue Internationale de l'Apprentissage en Ligne et de l'Enseignment À Distance* **2004**, *19*, 20–35.
20. Schwier, R.A.; Campbell, K.; Kenny, R. Instructional designers' observations about identity, communities of practice and change agency. *Australas. J. Educ. Technol.* **2004**, *20*. [CrossRef]
21. Schwier, R.A.; Campbell, K.; Kenny, R.F. Instructional designers' perceptions of their agency: Tales of change and community. In *Instructional Design: Case Studies in Communities of Practice*; IGI Global: Hershey, PA, USA, 2007; pp. 1–18.
22. Access the Capability Model. American Talent Development. Available online: https://www.td.org/capability-model/access (accessed on 5 August 2021).
23. Instructional Designer Competencies. Welcome to Ibstpi. 21 April 2016. Available online: https://ibstpi.org/instructional-design-competencies/ (accessed on 18 March 2022).
24. Martin, F.; Ritzhaupt, A.D. Standards and Competencies for Instructional Design and Technology Professionals [E-book]. In *Design for Learning*; 2020; p. 20. Available online: https://edtechbooks.org/id/standards_and_competencies (accessed on 21 April 2022).
25. ISTE Standards: Educators | ISTE. ISTE Standards: Educators. Available online: https://www.iste.org/standards/iste-standards-for-teachers (accessed on 10 March 2022).
26. Polanyi, M. *The Tacit Dimension*, 2009th ed.; The University of Chicago Press: Chicago, IL, USA, 1966.
27. Wagner, R.K.; Sternberg, R.J. Practical intelligence in real-world pursuits: The role of tacit knowledge. *J. Personal. Soc. Psychol.* **1985**, *49*, 436. [CrossRef]
28. McAdam, R.; Mason, B.; McCrory, J. Exploring the dichotomies within the tacit knowledge literature: Towards a process of tacit knowing in organizations. *J. Knowl. Manag.* **2007**, *11*, 43–59. [CrossRef]

29. Viale, R.; Pozzali, A. Cognitive Aspects of Tacit Knowledge and Cultural Diversity. *Model-Based Reason. Sci. Technol. Med.* **2007**, *64*, 229–244. [CrossRef]

30. Steiger, D.M.; Steiger, N.M. Instance-based cognitive mapping: A process for discovering a knowledge worker's tacit mental model. *Knowl. Manag. Res. Pract.* **2008**, *6*, 312–321. [CrossRef]

31. Johnson-Laird, P.; Byrne, R. Mental models website: A gentle introduction. *Recuper. El* **2000**, *22*. Available online: http://www.tcd.ie/Psychology/Ruth_Byrne/mental_models/index.html (accessed on 21 April 2022).

32. Bolade, S.; Sindakis, S. Micro-Foundation of Knowledge Creation Theory: Development of a Conceptual Framework Theory. *J. Knowl. Econ.* **2019**, *11*, 1556–1572. [CrossRef]

33. Chen, C.-H.; Saeedi, M. Building a Trust Model in the Online Market Place. *J. Internet Commer.* **2006**, *5*, 101–115. [CrossRef]

34. Dudek, A.; Patalas-Maliszewska, J. A Model of a Tacit Knowledge Transformation for the Service Department in a Manufacturing Company: A Case Study. *Found. Manag.* **2016**, *8*, 175–188. [CrossRef]

35. Jackson, T.W.; Tedmori, S.; Hinde, C.J.; Bani-Hani, A. The Boundaries of Natural Language Processing Techniques in Extracting Knowledge from Emails. *J. Emerg. Technol. Web Intell.* **2012**, *4*, 119–127. [CrossRef]

36. Mohanan, M.; Samuel, P. Software Requirement Elicitation Using Natural Language Processing. In *Innovations in Bio-Inspired Computing and Applications*; Springer: Cham, Switzerland, 2015; pp. 197–208. [CrossRef]

37. Satsangi, P. Automation of Tacit Knowledge Using Machine Learning. In Proceedings of the 2019 6th International Conference on Soft Computing & Machine Intelligence (ISCMI), Johannesburg, South Africa, 19–20 November 2019; pp. 35–39.

38. Stone, A.; Sawyer, P. Identifying tacit knowledge-based requirements. *IEE Proc. Softw.* **2006**, *153*, 211–218. [CrossRef]

39. Chambliss, M.J.; Garner, R. Do Adults Change their Minds after Reading Persuasive Text? *Writ. Commun.* **1996**, *13*, 291–313. [CrossRef]

40. Srivastav, A.; Singh, S. Proposed Model for Context Topic Identification of English and Hindi News Article Through LDA Approach with NLP Technique. *J. Inst. Eng. Ser. B* **2021**, *103*, 591–597. [CrossRef]

41. Shang, X.; Peng, Z.; Yuan, Q.; Khan, S.; Xie, L.; Fang, Y.; Vincent, S. DIANES: A DEI Audit Toolkit for News Sources. *arXiv* **2022**, arXiv:2203.11383.

42. Fung, Y.C.; Lee, L.K.; Chui, K.T.; Cheung, G.H.K.; Tang, C.H.; Wong, S.M. Sentiment Analysis and Summarization of Facebook Posts on News Media. In *Data Mining Approaches for Big Data and Sentiment Analysis in Social Media*; IGI Global: Hershey, PA, USA, 2022; pp. 142–154.

43. Singh, R.; Singh, S. Text Similarity Measures in News Articles by Vector Space Model Using NLP. *J. Inst. Eng. Ser. B* **2020**, *102*, 329–338. [CrossRef]

44. Yu, S.; Martino, G.D.S.; Nakov, P. Experiments in detecting persuasion techniques in the news. *arXiv* **2019**, arXiv:1911.06815.

45. Gao, N.; Touran, A.; Wang, Q. Mining and Visualizing Cost and Schedule Risks from News Articles with NLP and Network Analysis. *Constr. Res. Congr.* **2022**, 314–324. [CrossRef]

46. Jaidka, K.; Ceolin, A.; Singh, I.; Chhaya, N.; Ungar, L. WikiTalkEdit: A Dataset for modeling Editors' behaviors on Wikipedia. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; pp. 2191–2200.

47. Brugman, S. Introduction—Pandas-Profiling 3.0.0 Documentation. Pandas Profiling. 2021. Available online: https://pandas-profiling.github.io/pandas-profiling/docs/master/rtd/ (accessed on 6 August 2021).

48. Mueller, A. WordCloud for Python Documentation—Wordcloud 1.8.1 Documentation. WordCloud for Python. 2020. Available online: http://amueller.github.io/word_cloud/ (accessed on 6 March 2021).

49. Kaur, J.; Buttar, P.K. A systematic review on stopword removal algorithms. *Int. J. Future Revolut. Comput. Sci. Commun. Eng.* **2018**, *4*, 207–210.

50. Gerlach, M.; Shi, H.; Amaral, L.A.N. A universal information theoretic approach to the identification of stopwords. *Nat. Mach. Intell.* **2019**, *1*, 606–612. [CrossRef]

51. Natural Language Toolkit—NLTK 3.6.2 Documentation. Natural Language Processing Toolkit-NLTK. Available online: https://www.nltk.org/ (accessed on 6 August 2021).

52. Lorian, S. TextBlob: Simplified Text Processing—TextBlob 0.16.0 documentation. TextBlob: Simplified Text Processing. Available online: https://textblob.readthedocs.io/en/dev/ (accessed on 6 August 2021).

53. spaCy. Industrial-strength Natural Language Processing in Python. spaCy-Industrial-Strength Natural Language Processing. Available online: https://spacy.io/ (accessed on 6 August 2021).

54. Řehůřek, R. Gensim: Topic Modelling for Humans. 2009. Available online: https://radimrehurek.com/gensim/ (accessed on 6 August 2021).

55. Grootendorst, M. GitHub-MaartenGr/BERTopic: Leveraging BERT and c-TF-IDF to Create Easily Interpretable Topics. BERTopic. 2020. Available online: https://github.com/MaartenGr/BERTopic (accessed on 18 March 2022).

56. Reimers, N. Pretrained Models—Sentence-Transformers Documentation. Pre-Trained Models. 2021. Available online: https://www.sbert.net/docs/pretrained_models.html (accessed on 6 August 2021).

57. Chang, J.; Gerrish, S.; Wang, C.; Boyd-Graber, J.; Blei, D. Reading tea leaves: How humans interpret topic models. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; Volume 22.

58. Rosner, F.; Hinneburg, A.; Röder, M.; Nettling, M.; Both, A. Evaluating topic coherence measures. *arXiv* **2014**, arXiv:1403.6397.

59. Krotov, V.; Murray State University; Johnson, L.; Silva, L. University of Houston Legality and Ethics of Web Scraping. *Commun. Assoc. Inf. Syst.* **2020**, *47*, 539–563. [CrossRef]

60. Mancosu, M.; Vegetti, F. What You Can Scrape and What Is Right to Scrape: A Proposal for a Tool to Collect Public Facebook Data. *Soc. Media Soc.* **2020**, *6*. [CrossRef]

61. Catanese, S.A.; De Meo, P.; Ferrara, E.; Fiumara, G.; Provetti, A. Crawling facebook for social network analysis purposes. In Proceedings of the international conference on web intelligence, mining and semantics, Sogndal, Norway, 25–27 May 2011; pp. 1–8.

62. Washburn, A.N.; Hanson, B.E.; Motyl, M.; Skitka, L.J.; Yantis, C.; Wong, K.M.; Sun, J.; Prims, J.P.; Mueller, A.B.; Melton, Z.J.; et al. Why Do Some Psychology Researchers Resist Adopting Proposed Reforms to Research Practices? A Description of Researchers' Rationales. *Adv. Methods Pract. Psychol. Sci.* **2018**, *1*, 166–173. [CrossRef]

63. Acevedo, M.M.; Roque, G. Resisting the Deprofessionalization of Instructional Design. In *Optimizing Instructional Design Methods in Higher Education*; IGI Global: Hershey, PA, USA, 2019; pp. 9–26. [CrossRef]

64. Bean, C. *The Accidental Instructional Designer: Learning Design for the Digital Age*; American Society for Training and Development: Alexandria, VA, USA, 2014.

65. North, C.; Shortt, M.; Bowman, M.A.; Akinkuolie, B. How Instructional Design Is Operationalized in Various Industries for job-Seeking Learning Designers: Engaging the Talent Development Capability Model. *TechTrends* **2021**, *65*, 713–730. [CrossRef]