*Human Factors in Design, Engineering, and Computing*, Vol. 238, 2026, 2679–2688
https://doi.org/10.54941/ahfe1008492

# How Startups Communicate Their Positioning in the EdTech and Corporate Training Market: A Topic Modeling Analysis

**Manuel Kaiser and Tim Beichter**

Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering IAO, Nobelstraße 12, 70569 Stuttgart, Germany

## ABSTRACT

The workforce transformation in established industries and organizations is driven by dynamics in technologies and changes in demographics. Consequently, organizations' skill requirements are shifting, and upskilling is becoming an essential practice to address this transformation. At the same time, this environment opens new opportunities for entrepreneurs to develop business models and creative solutions to overcome these challenges. Accordingly, a new industry has recently developed that offers corporate training and EdTech solutions at the intersection of technology, learning, and upskilling. Although these startups play a critical role in workforce transformation from a practical standpoint, there is a gap in the research landscape. It is therefore also essential for stakeholders, such as potential customers and investors, to gain a better understanding of these startups. For this reason, this study explores 1,106 startups from the Crunchbase database using biterm topic modeling (BTM). In doing so, this analysis reveals five clusters, in which startups communicate their positioning: exam preparation, coaching and consulting, digital learning platforms, enterprise skill software and personalized learning apps. Furthermore, the results suggest that the startups in this sample focus primarily on customer benefits in their communicated positioning and in their Crunchbase descriptions. Finally, based on this exploration, the study then proposes a positioning map of startups in the EdTech and corporate training market.

**Keywords:** Corporate training, Crunchbase, EdTech, Startup, Topic modeling, Workforce transformation

## INTRODUCTION

Artificial intelligence (AI), sustainability and demographic change are drivers of new skill demands in organizations (World Economic Forum, 2025). Therefore, prior research discussed upskilling as a vital way to deal with this workforce transformation (e.g., Beichter & Kaiser, 2023; Daher-Armache et al., 2025; Jaiswal et al., 2022; Noe et al., 2014). In this context, upskilling means the training of individuals (e.g., employees) to update their existing skills or learn new skills (Beichter & Kaiser, 2023). Furthermore, a recent World Economic Forum report indicates that upskilling is becoming a strategic priority in organizations as a large majority of organizations (85 percent) anticipate prioritizing employee skill development as a central topic through 2030 (World Economic Forum, 2025). At the same time, the report estimates that the

Received August 14, 2026; Revised September 26, 2026; Accepted October 3, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

current skill set of the employees will be transformed during this period (World Economic Forum, 2025).

While established organizations are engaging with future skill requirements (Jaiswal et al., 2022), new job profiles (World Economic Forum, 2025), upskilling programs and workforce strategies (Noe et al., 2014), startups are utilizing this entrepreneurial opportunity to launch business models in the upskilling market (Fakih, 2025). Thus, the challenges faced by established market players are met here with innovative and scalable startup ideas (Glasner, 2022).

However, the landscape of these startups is less explored (Fakih, 2025), systematized and therefore fragmented, which makes it difficult for investors, customers and policymakers to navigate and evaluate the business models. More specifically, it remains unclear how startups communicate their positioning in this market and what value propositions they are communicating to their stakeholders. Positioning refers to the strategic choice of activities that distinguish an organization from its competitors (Porter, 1996), with the goal of permanently embedding a unique value proposition in the minds of the target audience (Ries & Trout, 2001). Building on this, positioning is part of organizational framing (Pahnke et al., 2026) through which startups use their stakeholder communication (Cornelissen & Werner, 2014). Therefore, this study focuses on communicated positioning rather than on strategic competitive positioning in general.

A better understanding of this is needed, as the market for upskilling solutions is generally heterogeneous and attracts different stakeholders (e.g. B2A, B2B, B2C) (Thunissen, 2023), features complex service structures (e.g., individual coaching, e-learning platforms) (Noe et al., 2014), and recent progress in the field of AI has further increased the dynamics of learning concepts (see for example Pereira et al., 2023).

To address this gap, this study explores the communicated positioning of 1,106 EdTech and corporate training startups by using Biterm Topic Modeling (BTM) based on their self-description in the Crunchbase database. In doing so, this study aims to answer the following research question: *How do startups in the EdTech and corporate training markets communicate their business positioning?*

## DATA COLLECTION AND ANALYSIS WITH TOPIC MODELING

To collect data and gain insight into communicated startup positioning, this study builds on startups identified from the online database Crunchbase (www.crunchbase.com). This is a main source of information on startups, containing details on the entire landscape of entrepreneurial activities (Dalle et al., 2017). Accordingly, many previous studies in the fields of management and entrepreneurship (Altmeier & Fisch, 2024; Kaiser & Kuckertz, 2025; Kim & Jung, 2026) have already used this database to develop research datasets. For example, Savin et al. (2023) use startup descriptions from Crunchbase to identify trends by applying topic modeling. Brown and Rocha (2020) develop a map of investments in startups during an exogenous shock by using financial data from Crunchbase.

Using the Crunchbase industry filter, this study identified startups that fall under "EdTech" or "Corporate Training". The EdTech industry was selected as a filter because startups in this sector also cooperate with corporate clients, thereby enabling employee training (Fakih, 2025). As the data collection process was conducted in January 2026, the startups must have had a founding date between July 1, 2022, and December 31, 2025. This approach is in line with Kuckertz and Hinderer (2026), as the Crunchbase data query considers startups that are a maximum of 3.5 years on the market. Furthermore, the startup entry in the database must include all necessary information about the company (e.g., description, headquarters) (Esposito et al., 2024) and startups must have their status listed as active in Crunchbase to be included in the final sample. By following these steps in Crunchbase, a dataset of 1,106 startups was then exported.

For the next step, the description field of the 1,106 startups was used to analyze the communicated positioning of these startups by using the R software. Since all startup descriptions are in English, the dataset was prepared for text analysis by removing standard stopwords (e.g., the, and, is). In addition to this, a topic-specific stopword list (Schofield et al., 2017) was created to reduce general terms (e.g., learning, training, platform) from further analysis. An initial analysis revealed that, based on the most frequent terms without an individual stopword list, the topic clusters are dominated by platforms and general terms around upskilling, such as training, education or learning. To address this issue, these terms were added to the individual stopword list to identify semantic differences in the descriptions at the content level. *The individual list includes the following keywords: platform, learning, education, training, company, provides, offers, solutions, services, students, online, edtech, educational, powered.*

After that, the preprocessed data were used in RStudio for analysis with biterm topic modeling (BTM) from Yan et al. (2013). The approach of using BTM (see also Wijffels, 2025) for startup descriptions was also used by Kuckertz and Hinderer (2026), for example, to examine startup descriptions of the twin transformation. Given that startup descriptions from Crunchbase are brief texts, BTM is well suited as it specializes in such formats (Yan et al., 2013). Furthermore, computer-based text analysis with topic modeling is suitable to analyze organizational positioning (e.g., Haans, 2019).

The initial descriptive analysis shows that the 1,106 startups are in 88 countries, and most startups are headquartered in the United States (n = 358), India (n = 191) and the United Kingdom (n = 69). Furthermore, the average age of a startup is 26 months, while the oldest is 41 months old and the youngest is seven days old. On the financial side, 286 startups (26 percent) received external funding from investors with a total funding amount of 868,274,354.00 USD.

As a result of the biterm analysis, BTM identified five topic clusters (Table 1) from the data sample: exam preparation, coaching and consulting, digital learning platforms, enterprise skill software, personalized learning apps. In previous studies using topic modeling approaches, the topic selection (k) has been noted as a challenge (Savin et al., 2023). To address this, our study uses the following criteria: First, the distinction between the

clusters was taken into account. For this purpose, different values of "k" (2,3,4,5,6,7,8) were generated with BTM (1,000 iterations; Gibbs sampling algorithm), and the respective topics were analyzed for their distinctiveness. Regarding distinctiveness, the main question in this step was to what extent the individual keywords within each cluster were unique, such that separate topic clusters could be formed. Second, for each k the coherence value was calculated to select the model with the highest coherence score. In this step, the values ranged from –250.37 (k = 5) to –297.52 (k = 4). Third, a prevalence of 10 percent per topic was used as the threshold level (Kuckertz & Hinderer, 2026) to ensure that each topic represents a substantial share of the sample rather than a marginal phenomenon. Overall, this resulted in k = 5 with coherence score of –250.37 ($\beta = 0.01$). The five clusters were then subjected to a qualitative assessment and given the following names: *exam preparation, coaching and consulting, digital learning platforms, enterprise skill software, personalized learning apps*.

**Table 1:** Results of biterm analysis.

| Topic | Tokens (Probability) | Topic Prevalence |
|---|---|---|
| Exam preparation | study (0.019030), exams (0.013663), exam (0.012004), personalized (0.010638), abroad (0.009467), career (0.008101), tools (0.008003), college (0.007710), preparation (0.007710), guidance (0.007613) | 0.1572482 |
| Coaching and consulting | coaching (0.014045), consulting (0.013143), development (0.012092), digital (0.010365), business (0.009614), career (0.009088), support (0.008863), firm (0.008412), tools (0.008337), app (0.008037) | 0.2043612 |
| Digital learning platforms | tech (0.014976), courses (0.014405), skills (0.013122), data (0.012123), digital (0.010982), development (0.008629), support (0.008629), technology (0.008558), professionals (0.008201), global (0.007773) | 0.2152334 |
| Enterprise skill software | technology (0.013231), skills (0.011527), digital (0.009923), job (0.009422), based (0.008821), tools (0.008420), management (0.008019), saas (0.008019), software (0.007618), skill (0.007017) | 0.1531020 |
| Personalized learning apps | personalized (0.016655), teachers (0.013017), content (0.012051), helps (0.012051), tools (0.011994), app (0.011255), language (0.011027), designed (0.010345), time (0.009947), interactive (0.009436) | 0.2700553 |

## DISCUSSION OF THE RESULTS

The analysis of the five topics shows that they differ depending on whether the startups target individuals or organizations. For example, there are solutions designed to support individual learners, as well as solutions that primarily support skill management in organizations through software applications. Furthermore, the results suggest that while people-centered approaches (e.g., coaching, consulting) do exist, they are primarily extended through digital channels (e.g., digital platforms, apps). The "exam preparation" cluster suggests that this is not only a traditional corporate training program, but rather one aimed at students (e.g., high school, college). One possible explanation lies in the diversity and breadth of the EdTech industry. A Crunchbase report shows that EdTech solutions cater to a wide range of audiences, from K-12 and college students to established companies (Glasner, 2025).

An exploratory observation of this analysis suggests that AI-related wording does not dominate the startups' Crunchbase descriptions and thus in their communicated positioning in this study. While the sample covers the period from 2022 to 2025, it is precisely during this time that generative AI (with the launch of ChatGPT in November 2022) gained momentum. Discussions surrounding AI have also intensified in the context of upskilling (e.g., Randazzo et al., 2025) and beyond that, also in the area of business models (e.g., Jobstreibizer et al., 2025). However, the identified five topics still show a stronger connection to digitalization (e.g., app, digital, software, technology) in their Crunchbase descriptions than to AI topics. This is even more surprising when looking in detail at the industries in which the startups are listed on Crunchbase. Over 300 startups in this study are listed under AI category in the Crunchbase database. It should also be noted that the analysis is limited to the 10 key terms per cluster. If this limitation is expanded to 20 words per cluster, it becomes apparent in the "Personalized learning apps" cluster that intelligence (probability: 0.007958080) and artificial (probability: 0.007048650) play a role here. In sum, this could indicate that in communicated positioning, differentiation is framed less around AI terms and more about customer value (e.g., personalized, interactive). And particularly in the case of personalized learning, this can certainly be AI-supported (e.g., Pereira et al., 2023); however, the value should lie in the personalization itself, with the technology serving as the tool for implementing it (e.g., personalized apps).

In the five clusters, it is noticeable that the descriptions do use business-related terms (e.g., career, professionals, job), but do not include further details in the communicated positioning. For example, there are different job roles within companies (World Economic Forum, 2025), ranging from production to office work and management roles. In this respect, it is not apparent in the identified clusters whether the startups focus on specific target groups here and reflect this in their communicated positioning. Furthermore, training needs can differ not only at the level of job roles, but also in terms of varying skill demands depending on the industry. Here, too, the results suggest that industry-specific focuses are not incorporated into the description as a key distinguishing feature.

The Crunchbase descriptions of these startups also lend themselves to analysis from a business model perspective (Kuckertz & Hinderer, 2026). The results show that the Software-as-a-Service (SaaS) model appears in the enterprise skill software cluster. This finding suggests that some startups communicate their software-based offerings to their customers.

Building conceptually on these five clusters, Figure 1 proposes a positioning map to illustrate the explorative findings. To this end, the authors suggest identifying the different target groups (e.g., firm, career, professionals) on the one hand with "individual vs. organization" and mapping out the focus of the value proposition (e.g., coaching, personalized, app) on the other with "learning vs. technology". Overall, the map should not be understood as an empirical classification, but rather as a tool to highlight key conceptual communication positioning characteristics of startups in the EdTech and corporate training sectors.

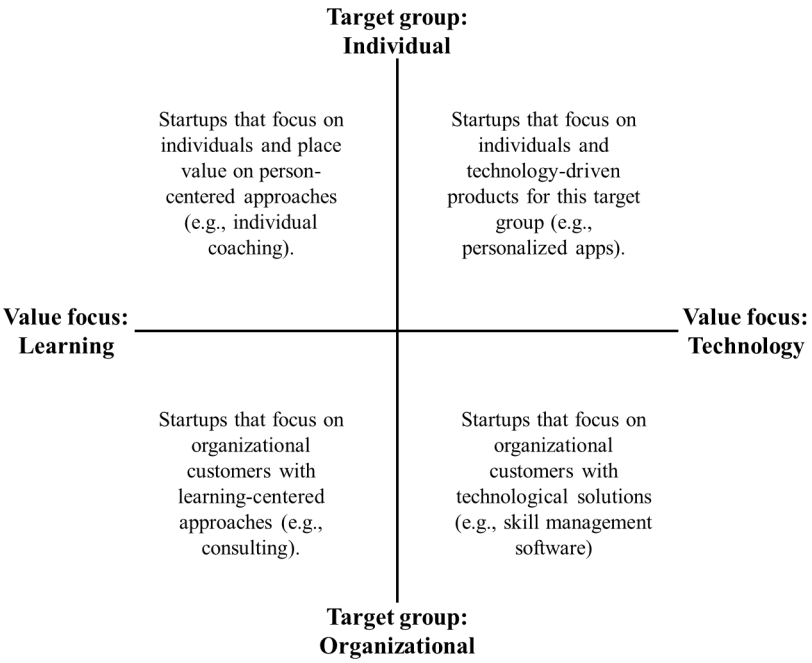



**Figure 1:** Conceptual positioning map of startups in the EdTech and corporate training market.

## FUTURE RESEARCH AVENUES, PRACTICAL IMPLICATIONS AND LIMITATIONS

While literature in both research and practice contains established concepts and models for determining a company's positioning from a strategy (Porter, 1996; Treacy & Wiersema, 1993) or marketing-communication perspective (Ries & Trout, 2001), the transfer to entrepreneurship research is still in its infancy. This study bridges that gap by analyzing startup descriptions from Crunchbase using topic modeling. Further studies could

build on this by analyzing additional industries or incorporating the individual communicated positioning of founders. In particular, the field of entrepreneurs' communication (Kaiser & Kuckertz, 2024) opens further research opportunities here to classify communicative positioning based on the language used. More specifically, the role of AI in positioning and communication strategies could be further explored to determine whether AI is being integrated more deeply into communication across different channels (e.g., social media, websites, press interviews). Furthermore, the perspective of investors (e.g., venture capitalists, business angels) could be included here to explore the importance they place on AI in the communicated positioning of startups and the role they see in evaluating startups.

The sample for this study includes a total of 1,106 startups, headquartered in 88 countries. This also includes startups that offer their services locally, as well as startups that target international markets. At the same time, a recent analysis by the World Economic Forum (2025) shows that requirements for future skills and upskilling vary between countries, and that different needs may arise across nations. This opens up a research opportunity in the field of EdTech and corporate training startups. Future studies could, for example, analyze the business models of these startups and conduct an international comparative analysis.

The five identified topic clusters highlight the heterogeneity of the upskilling market and its fragmentation into diverse offerings and target groups. This suggests that there are startups with solutions tailored to different target groups and requirements. For established companies as potential customers (B2B), this exploratory mapping could serve as orientation to evaluate startups in the upskilling market. Furthermore, the findings and topic clusters presented here can help them clarify their own requirements.

Since learning is also an area of investment for investors in the startup ecosystem (Antonenko et al., 2014; Avelar, 2025), this study provides these stakeholders with an overview of a fragmented market. Accordingly, investors can also use these results as a map and for navigation (e.g., regarding geographical focus and topic clusters in the EdTech and corporate training sectors).

Besides the implications for researchers and practitioners, however, this study also has limitations. (1) In classifying the communicated positioning, it is important to note that this is based primarily on Crunchbase data, and no additional data was used to supplement the communicated positioning. This also means that only startups listed in the database at the time of the Crunchbase query were included. (2) Furthermore, the study focused exclusively on startups operating in the EdTech and corporate training sectors. If a startup in the field of human resource management (e.g., employee development) develops solutions but is not classified under these two Crunchbase industries, it is not included in this study. (3) The BTM analysis was the main approach to form the five clusters. Accordingly, this study does not include an in-depth analysis or provide links to additional information about the startups (e.g., clusters combined with industries or

financing information). (4) With regard to the description texts used, it cannot be ruled out that they were either optimized by marketing experts or written with (generative) AI tools. Furthermore, these texts reflect the communicated positioning of startups rather than verified product or business model features. (5) Using the conceptual map, this study also proposes a possible classification; however, this is a qualitative classification that has not yet been empirically verified in any greater depth.

## CONCLUSION

Startups in the field of EdTech and corporate training play an important role in workforce transformation, both for organizations and for employees who use these offerings. While research on EdTech startups has already produced earlier publications (Antonenko et al., 2014; Fakih, 2025), the transfer to established companies and especially the applications in the area of upskilling still offer considerable potential for future research. In conclusion, it should be noted that this is primarily an exploratory study, the results of which will be further examined in subsequent empirical studies.

## REFERENCES

Altmeier, A., & Fisch, C. (2024). The Relationship Between Business Angels' Psychological Attributes and Investments in Successful Portfolio Ventures: An Empirical Investigation Using Twitter Data. British Journal of Management, 35(2), 629–643. https://doi.org/10.1111/1467-8551.12783

Antonenko, P. D., Lee, B. R., & Kleinheksel, A. J. (2014). Trends in the crowdfunding of educational technology startups. TechTrends, 58(6), 36–41. https://doi.org/10.1007/s11528-014-0801-2

Avelar, M. (2025). Mapping the digital education landscape: Stakeholders and networks of governance in Brazil. International Journal of Educational Development, 112, 103192. https://doi.org/10.1016/j.ijedudev.2024.103192

Beichter, T., & Kaiser, M. (2023). The Future of Upskilling: Human- and Technology-Centered. In N. Callaos, N. Callaos, S. Hashimoto, N. Lace, B. Sánchez, & M. Savoie (Eds.), Proceedings of the International Multi-Conference on Complexity, Informatics and Cybernetics, Proceedings of the 14th International Multi-Conference on Complexity, Informatics and Cybernetics: IMCIC 2023 (pp. 190–193). International Institute of Informatics and CyberneticsWinter Garden, Florida, United States. https://doi.org/10.54808/IMCIC2023.01.190

Brown, R., & Rocha, A. (2020). Entrepreneurial uncertainty during the Covid-19 crisis: Mapping the temporal dynamics of entrepreneurial finance. Journal of Business Venturing Insights, 14, e00174. https://doi.org/10.1016/j.jbvi.2020.e00174

Cornelissen, J. P., & Werner, M. D. (2014). Putting Framing in Perspective: A Review of Framing and Frame Analysis across the Management and Organizational Literature. Academy of Management Annals, 8(1), 181–235. https://doi.org/10.1080/19416520.2014.875669

Daher-Armache, G., Armache, J., & Ismail, H. N. (2025). Leadership development in U.S. Higher education: Strategies for lifelong learning and upskilling. Journal of Innovation & Knowledge, 10(4), 100754. https://doi.org/10.1016/j.jik.2025.100754

Dalle, J.-M., den Besten, M., & Menon, C. (2017). Using Crunchbase for economic and managerial research. OECD Science, Technology and Industry Working Papers. https://doi.org/10.1787/6c418d60-en

Esposito, C. D., Szatmari, B., Sitruk, J. M. C., & Wijnberg, N. M. (2024). Getting off to a good start: emerging academic fields and early-stage equity financing. Small Business Economics, 62(4), 1591–1613. https://doi.org/10.1007/s11187-023-00816-9

Fakih, A. (2025). Analysing the Contributions of Edtech Startups to the Corporate Sector. Journal of Contemporary Business Research, 1(1), 34–46. https://doi.org/10.1177/3049513X251319635

Glasner, J. (2022). 'Great Resignation' Drives Billions In VC Dollars To Upskilling Startups. https://news.crunchbase.com/startups/worker-upskilling-training-reskilling-startups-vc-funding-guild-education/ (Last accessed: September 6, 2026)

Glasner, J. (2025). Edtech-Specific Startup Funding Stays Low. https://news.crunchbase.com/venture/edtech-funding-stays-low/ (Last accessed: September 6, 2026)

Haans, R. F. J. (2019). What's the value of being different when everyone is? The effects of distinctiveness on performance in homogeneous versus heterogeneous categories. Strategic Management Journal, 40(1), 3–27. https://doi.org/10.1002/smj.2978

Jaiswal, A., Arun, C. J., & Varma, A. (2022). Rebooting employees: upskilling for artificial intelligence in multinational corporations. The International Journal of Human Resource Management, 33(6), 1179–1208. https://doi.org/10.1080/09585192.2021.1891114

Jobstreibizer, J., Beliaeva, T., Ferasso, M., Kraus, S., & Kallmuenzer, A. (2025). The impact of artificial intelligence on business models: a bibliometric-systematic literature review. Management Decision, 63(13), 372–396. https://doi.org/10.1108/MD-10-2024-2309

Kaiser, M., & Kuckertz, A. (2024). Bibliometrically mapping the research field of entrepreneurial communication: where we stand and where we need to go. Management Review Quarterly, 74(4), 2087–2120. https://doi.org/10.1007/s11301-023-00355-3

Kaiser, M., & Kuckertz, A. (2025). Emotions and entrepreneurial finance: Analysis of venture capitalists' and business angels' digital footprints on Twitter. International Entrepreneurship and Management Journal, 21(1). https://doi.org/10.1007/s11365-024-01018-9

Kim, Y., & Jung, H. J. (2026). When do ventures break path dependence? Evidence from financial and technological success of serial entrepreneurs. Technovation, 149, 103367. https://doi.org/10.1016/j.technovation.2025.103367

Kuckertz, A., & Hinderer, S. (2026). Selective coupling or genuine integration? Startups in the twin transformation. Sustainable Futures, 11, 101675. https://doi.org/10.1016/j.sftr.2026.101675

Noe, R. A., Clarke, A. D., & Klein, H. J. (2014). Learning in the Twenty-First-Century Workplace. Annual Review of Organizational Psychology and Organizational Behavior, 1(1), 245–275. https://doi.org/10.1146/annurev-orgpsych-031413-091321

Pahnke, E. C., Zuzul, T., & Howard, M. (2026). Strategic positioning and accelerating regulatory clearance for new ventures. Strategic Management Journal, 47(6), 1730–1763. https://doi.org/10.1002/smj.70072

Pereira, D. S., Falcão, F., Costa, L., Lunn, B. S., Pêgo, J. M., & Costa, P. (2023). Here's to the future: Conversational agents in higher education- a scoping review. International Journal of Educational Research, 122, 102233. https://doi.org/10.1016/j.ijer.2023.102233

Porter, M. E. (1996). What is strategy? Harvard Business Review, 74(6), 61–78.
Randazzo, S., Lifshitz-Assaf, H., Kellogg, K., Dell'Acqua, F., Mollick, E. R., Candelon, F., & Lakhani, K. R. (2025). Cyborgs, Centaurs and Self Automators: The Three Modes of Human-GenAI Knowledge Work and Their Implications for Skilling and the Future of Expertise. Harvard Business School (Working Paper 26-036).
Ries, A., & Trout, J. (2001). Positioning: The Battle for Your Mind. McGraw-Hill Education.
Savin, I., Chukavina, K., & Pushkarev, A. (2023). Topic-based classification and identification of global trends for startup companies. Small Business Economics, 60(2), 659–689. https://doi.org/10.1007/s11187-022-00609-6
Schofield, A., Magnusson, M., & Mimno, D. (2017). Pulling Out the Stops: Rethinking Stopword Removal for Topic Models. In M. Lapata, P. Blunsom, & A. Koller (Eds.), Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers (pp. 432–436). Association for Computational Linguistics.
Thunissen, M. (2023). A stakeholder approach to Lifelong Learning. European Journal of University Lifelong Learning, 7(2), 5–12. https://doi.org/10.53807/0702sA0x
Treacy, M., & Wiersema, F. (1993). Customer intimacy and other value disciplines. Harvard Business Review, 71(1), 84–93.
Wijffels, J. (2025). BTM: Biterm Topic Models for Short Text. https://cran.r-project.org/web/packages/BTM/index.html (Last accessed: September 6, 2026)
World Economic Forum. (2025). Future of Jobs Report. https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf (Last accessed: September 6, 2026)
Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. In D. Schwabe, V. Almeida, H. Glaser, R. Baeza-Yates, & S. Moon (Eds.), Proceedings of the 22nd international conference on World Wide Web (pp. 1445–1456). ACM. https://doi.org/10.1145/2488388.2488514