Data Systems Education in Finland: Aligning Higher Education with Industry Expectations

Nea Peltola
nea.peltola@tuni.fi
Tampere University
Tampere, Finland

Toni Taipalus
toni.taipalus@tuni.fi
Tampere University
Tampere, Finland

Abstract

Courses on data systems in higher education are usually designed based on curriculum guidelines, textbooks, the instructor's own understanding of the subject, or a mix of these sources. As a result, the course contents may not always align with current industry standards, either in terms of theoretical foundations or the technical tools used to teach data systems concepts. In this mixed-methods study, we investigate the technical and non-technical skills needed in data systems industry, with the goal of bridging the gap of data systems education and the data systems industry. Our results show that SQL, data modeling and data pipelines, as well as cloud computing and data warehouses are the most sought after skills in Finnish industry. Our results are applicable in developing data systems syllabi and curricula to better meet industry expectations for future data professionals.

CCS Concepts

• **Applied computing → Education**; Psychology; Sociology; • **Information systems → Data management systems**.

Keywords

database, higher education, data literacy, misinformation, industry, curriculum, data system

ACM Reference Format:
Nea Peltola and Toni Taipalus. 2026. Data Systems Education in Finland: Aligning Higher Education with Industry Expectations. In *Proceedings of the 31st ACM Conference on Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2026), July 10–15, 2026, Madrid, Spain.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3803400.3809334

1 Introduction

Data-related skills are more and more relevant in the industry (i.e., *not academia*), as data-driven technologies and societal changes in the information landscape become increasingly common. It follows that data-related skills such as Structured Query Language (SQL) [6], data science [32], and data literacy [14, 22] are in high demand in today's job markets. However, several studies over the years have highlighted the challenges and discrepancies of industry expectations and higher education [3, 25]. In other words, higher education courses do not necessarily cater for the needs of data systems industry, as higher education courses are often based on rarely-updated curriculum guidelines, textbooks, and teacher opinions [25]. It also remains unclear how well different international curriculum guidelines particularize to different countries.

In this study, we investigated industry needs for data systems skills in Finland. In particular, we explored which technical, data-related skills (e.g., data modeling, database optimization) are valued in industry positions, as well as how societal, data-related skills (e.g., data literacy) should be taught in higher education, and why such societal skills are important in the first place.

Our results detail which technical skills are valued in the data systems industry (e.g., SQL, data modeling), why each of these skills are deemed necessary, and which technologies (e.g., Snowflake, DuckDB) are used. Furthermore, we provide analysis on which adjacent skills (e.g., working with stakeholders, collaboration) are valued and why. Finally, we provide insights on *societal* data skills (e.g., identifying mis- and dis-information), why they are deemed necessary in today's information landscape, and how they can be taught in practice.

The rest of this study is structured as follows. In the next Section, we discuss prior work on data systems skills needed in the industry, as well as societal data skills. In Section 3, we describe our survey instrument, interviews, and data analysis. Section 4 contains the results, Section 5 the practical implications and threats to validity of this study, and Section 6 concludes the study.

2 Theoretical background

2.1 Industry needs versus curricula

There is a recognized gap between industry expectations and higher education, and this misalignment is especially prominent in fast-evolving fields such as data science and big data [21, 26]. While higher education often values theoretical knowledge, students arguably study for their future work life, where many roles involve highly practical tasks. It has been argued that higher education must maintain the practical utility of courses and continuously adapt to job market demands [21].

Data systems is a core area of computing, encompassing concepts such as database design, data engineering, transaction processing, and data analytics [25]. The choices of data systems skills included in computing curricula are often guided by textbooks, teacher opinions, and various curriculum guidelines [e.g., 11, 12, 23]. As there are various expert opinions on different levels of abstraction on what *data systems* effectively means in education, a previous study analyzed thirteen data systems textbooks, six curriculum guidelines, and three national guidelines, and formulated a framework of 24 *technical* data systems skills (i.e., knowledge areas) such as SQL, data warehousing, and database management system (DBMS) components [25]. These skills are listed in Section 4.1. This framework

ITiCSE 2026, Madrid, Spain

ACM ISBN 979-8-4007-2634-7/2026/07
https://doi.org/10.1145/3803400.3809334

guides the data collection in our study, as it provides a consistent set of skill definitions to compare with industry expectations. There has also been an effort in mapping data systems skills to closely related fields (e.g., data science and computer engineering) in education [7, 20].

2.2 Industry expectations for data systems skills

Prior studies show strong demand for skills like SQL [cf. e.g., 30] and database design, but also for newer areas such as NoSQL databases [cf. e.g., 31], data visualization, and data warehousing, which are, however, seldom taught in courses [25]. Similarly, job-market analyses in big data roles report that employers prioritize competencies in areas like distributed data processing, cloud data platforms, and business intelligence [21], as well as data analysis and data warehousing [13]. Job-market analysis has also highlighted the need for big data frameworks, big data processes, and NoSQL databases, alongside Java/Python/Scala-centric programming, but also *soft skills* such as teamwork and communication [17]. Additionally, Python, R, and Tableau skills have been seen important in data analyst roles [8].

Finnish higher education follows the so-called *dual education* model, where Finnish universities focus on education from the perspectives of scientific research and producing new knowledge, whereas universities of applied science focus on more vocational aspects, and on research and development. For the latter, keeping curricula up-to-date with fast-moving information technology industry is even more crucial. To the best of our knowledge, efforts to align Finnish data systems education with industry expectations are not reported in scientific literature.

2.3 Societal skills

In the modern information landscape, we recognize *societal* data systems skills in addition to *technical* (e.g., database optimization), as well as *soft* skills (e.g., teamwork). Societal skills are closely related to data systems, particularly to data interpretation, cleaning, privacy, and organizational and personal decision-making. Managing data in such an objective and ethical manner has been recognized as a valued skill in industry [5].

Based on a relatively cursory search of scientific literature, we identified six societal data systems skills: *critical evaluation of data quality* [19], *identifying mis- and disinformation* [15, 16, 28], *interpreting and comparing data* [2], *data-driven decision making* [1], *ethical use of data* [16], and *data privacy and sharing* [16, 28]. It is worth noting that there is a level of overlap among these skills, but they both give structure to subsequent data collection and analysis, and paint a picture of what is meant by societal skills on a high level of abstraction.

These types of societal data skills arguably complement technical data skills such as data cleaning and aggregation, data visualization, performance monitoring, and A/B testing. Studies on strengthening data literacy have shown positive effects in identifying misinformation [29], understanding data visualizations [4], and making data-driven decisions [27].

Table 1: Questions in the structured technical and societal interviews

#	Question
T1	What is your current role and background in data-related tasks?
T2	For two to six skills the participant had evaluated critical or important in the survey instrument, repeated for each area:
T2.1	Why is it important to have this skill, and where is it utilized in practice?
T2.2	Which tools and technologies are utilized with this skill?
T2.3	Why those particular tools and technologies were chosen?
T3	Is there a skill in the survey instrument that is unnecessary?
T4	What kind of skills should data systems education focus on?
T5	What new technologies and subjects related to data systems have arisen lately?
T6	What kind of non-technical skills are required in the industry?
S1	What are the essential non-technical data skills that everyone should know?
S2	Repeated for each skill [Critical evaluation of data quality, Identifying mis- and disinformation, Interpreting and comparing data, Data-driven decision-making, Ethical use of data, Data privacy and sharing]: How could this skill be taught in higher education?
S3	Are there any other skills that should be included in this set of societal data skills?

3 Methodology

3.1 Survey instrument

The survey consisted of three background questions regarding the field of industry, company size, and working location within Finland, four-point Likert questions regarding 24 data systems skills (cf. Fig. 1, *"Evaluate the necessity of each skill for entry-level positions requiring database knowledge"*), and one free-form text field in which the participant could fill in required skills not mentioned in the Likert questions. The 24 skills used in this study were recognized in an earlier data systems curriculum analysis [25]. Before the questions proper, a full privacy policy was shown to potential participants, detailing how their data would be handled. Participation was voluntary, and carried no immediate positive or negative effects for the participants. The survey targeted professionals working with data systems in industry, without limiting participation to specific job roles. We distributed the survey to known contacts, as well as through posts on social media. In total, we received responses from participants in various industries (N = 117), mainly from IT. The data were collected during spring 2025.

3.2 Interviews

We conducted structured interviews (N = 34) to gain more insights into both technical and societal aspects of data systems education. Each participant only participated in either a *technical* (n = 25) or a *societal* interview (n = 9). The participants of the societal interviews were researchers familiar with the subject, recruited from Finnish universities. Participants of the technical interviews had also responded to the survey. These questions are listed in Table 1.

The interviews were held both remotely and in-person. The interviews were conducted by the first author. Due to the relatively straightforward nature of the interview questions, the first author

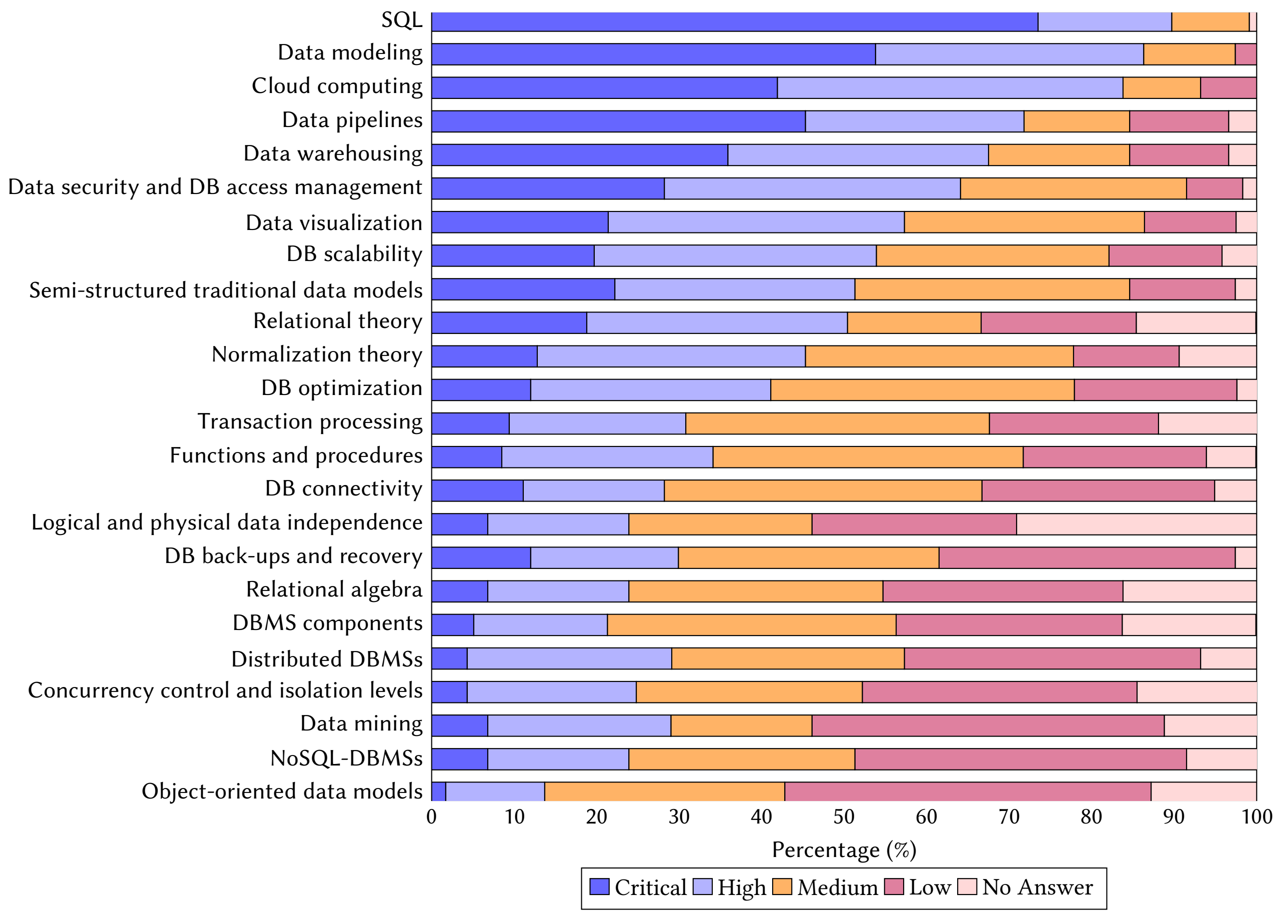

Figure 1: Proportions of different data system skills, ordered according to the Likert mean of how necessary a skill was deemed by the survey respondents

took notes during the interviews, and the interviews were not recorded. Due to the privacy policy agreed by the respondents, we do not show direct quotations from the interviews, but summaries of recurring themes. The average lengths of the interviews were 30 minutes (technical interviews), and 25 minutes (societal interviews). Because all individuals took part voluntarily with informed consent, none were minors, and the activity posed no safety risks, obtaining ethical approval for data collection was not required under the national research integrity board's guidelines.

3.3 Analyses

We report the technical data systems skills as descriptive statistics in Section 4. Additionally, we tested for differences between *(1)* different company sizes (75 respondents from large companies versus 42 respondents from companies of other sizes) and *(2)* different industries (72 respondents from IT versus 45 respondents from other industries) using Welch's t-test due to unequal variances. Finally, we utilized conventional content analysis [18] in finding the technical and non-technical skills deemed essential in the free-text responses of the survey.

We analyzed the interview notes using conventional content analysis [18]. The first and second author individually coded 10% of the interviews, one interview question at the time. Next, the authors compared their codes. As the inter-rater agreement was deemed high, possibly due to the rather straightforward research questions, and the first author continued to code the remaining data. After coding, the authors convened to summarize the raw codes into categories.

4 Results

4.1 Necessary skills in data systems industry

Survey answers regarding the 24 data systems skills are visualized in Fig. 1 according to how necessary a skill was deemed in the industry by the survey respondents. The most necessary skills were SQL, data modeling, cloud computing, data pipelines, and data warehousing, respectively.

4.2 The effects of company size and industry

We chose an alpha level of .05 for all the statistical tests, and adjusted with Bonferroni correction to account for the number of 24 data

Table 2: Data systems skills, why they are deemed necessary in industry, and necessary technologies mentioned by the interviewees; *object-oriented data models* and *database distribution* received no arguments for being necessary

Skill	Why the skill is deemed necessary	Technologies
SQL	Prominent in the industry; An important skill also in the future; Still in the background in many systems; Prominent tool in reporting; Mature technology; The first fall-back tool when something goes wrong; Computationally fast; Understanding SQL helps in optimization; Employers expect knowledge of SQL	PostgreSQL, Oracle DB, MySQL, open source systems, small embedded databases, AlloyDB, DuckDB
Data modeling	Prominent in the industry; Difficult to do much without this skill, especially in data warehouses; Formalization of the client's needs into a database structure and sharing this information with other people; Especially important when designing changes to the database structure; The same database may be modeled in different ways; Understanding in how design choices affect storage requirements and execution time	An agreed upon notation, pen and paper
Cloud computing	Prominent in the industry; New systems are often required to be in the cloud; Alternatives are inflexible; Some domains have too much data for alternative solutions; Seen as cost-effective	AWS, Google Cloud, Azure, Snowflake, Amazon Redshift, European cloud solutions
Data pipelines	Prominent in the industry; Used for ensuring data uniformity between different data sources; Understanding and ensuring the reliability of data lineage; It is a modern technique	Azure Data Factory, Snowflake, dbt, Databricks, Data Fabric, Python, Java
Data warehousing	Needed for AI solutions; Closely related to reporting which is prominent in industry; The amount of teaching is limited given how necessary the skill is in industry	Azure SQL, AWS, Snowflake, Data Vault 2.0, Terraform, Agile Data Engine, PostgreSQL, SQL Server, Databricks, dbt
Data security and DB access management	In the initial setup; Reliability and trustworthiness from the public and client's perspectives; Data privacy is crucial; The amount of teaching is limited given how necessary the skill is in industry	
Data visualization	Important for utilizing data in decision making; Presenting data in an understandable form	PowerBI, Tableau, ClickView, Excel, Python libraries
DB scalability	Understand the differences between development and production environments and different database sizes; In the cloud, scalability directly affects costs; In the industry, database sizes are different to educational contexts; Ensuring business continuity	
Semi-structured traditional data models	Important in combining data from various sources, and sharing outbound data in required formats; Used in reporting	XML, JSON, R, Python
Relational theory	Theoretical principles for data modeling; Relational databases are prominent in the industry	
Normalization theory	Theoretical principles for data modeling; Used in formalizing domain requirements	
DB optimization	Affects cost efficiency; Ensuring business continuity; More optimized is greener; Basic knowledge of optimization is important, because then the developer understands to leave much of the optimization to the DBMS	Query execution plans
Transaction processing	Understanding basic principles for optimization and anomaly prevention; Crucial in multi-user environments	
Functions and procedures	Performance considerations	
DB connectivity	Understanding the system from a more holistic perspective	JDBC, ORMs
Logical and physical data independence	Understanding technology-independent principles	
DB back-ups and recovery	Trustworthiness from the public and client's perspectives; Ensuring business continuity; Used in differentiating critical and non-critical data	
Relational algebra	Theoretical principles for querying	
DBMS components	Important in processing large databases	
Concurrency control and isolation levels	Crucial in multi-user environments; Helps understanding complex systems	
Data mining	Prominent with AI	Python, R
NoSQL-DBMSs	Prominent in the industry	MongoDB, some features of PostgreSQL, Firebase, key-value stores

systems skills. That is, we chose an alpha level of $\alpha = \frac{0.05}{24} = .003125$ for the comparisons.

We conducted Welch's independent samples t-tests to examine whether the data systems skill valued differed between large companies and companies of other sizes. There were no significant differences in mean scores between the two groups in any of the 24 skills, p-values ranging approximately from .05 to .95.

We conducted Welch's independent samples t-tests to examine whether the data systems skill valued differed between IT companies and companies of other industries. There were no significant differences in mean scores between the two groups in any of the 24 skills, p-values ranging approximately from .10 to 1.00.

4.3 Why some skills are deemed necessary

In the technical interviews, we asked the participants why they deemed some of the skills necessary and some unnecessary. We summarized the answers for the former question to Table 2.

For the latter question, the participants voiced a common reason regarding all the skills: they have not needed that skill personally, or seen others need it recently. In fact, we were able to find only two insights from the analysis of this question. First, *cloud computing* was seen unnecessary by some of the participants due to both small-scale applications which do not necessarily benefit much from using a cloud environment, and the fact that cloud environments pose data privacy issues, and are unnecessary to teach in higher education. Second, and quite adversarially, *database back-ups*, *database scalability*, and *database optimization* were seen unnecessary due to cloud environments offering much of these features as a service.

4.4 Additional necessary skills

In addition to the 24 skills identified for the survey (cf. Fig. 1), both the survey participants and the technical interviewees also voiced other skills necessary in industry. We summarized the groups of skills which were mentioned at least by six participants below.

From a technical perspective, several participants emphasized that it is crucial to be familiar with *modern database technologies* such as data vaults, data lakes, the Parquet file format, and DuckDB; *non-relational data models* such as vector databases, time-series databases, and graph databases; *holistic system perspectives*, i.e., understanding the adjacent technologies, people, and processes, infrastructure, and enterprise architecture, and that the database does not exist in isolation; *performance*, i.e., understanding green aspects of databases and the implications of using languages such as Python in data pipelines, and the effects of design choices on computing resources; *software development frameworks* such as Scrum and Kanban and continuous development, and how they affect data; *testing* such as test data generation, software testing and its implications to the database, and the role of a myriad of data sources on testing; and *data governance* such as processes and tools for protecting data, auditing data usage, and ensuring data quality. Additionally, several participants speculated that the advancements in artificial intelligence will affect data management in some way.

From a non-technical perspective, the participants highlighted *domain knowledge*: understanding the business environment and business processes, related terms, and the scale of the domain; *working with stakeholders*: understanding and communicating the needs of the client and other related parties; *transforming business rules into data*: knowing which business rules can dictate the use of different database paradigms, and which business rules are relevant from a database point-of-view; *interaction skills*: communicating with others, expressing oneself clearly, diplomacy, language skills in international teams, and communication with non-technical audiences; *collaboration*: working in a team, project work, and social skills; and *self-governance*: being independent when the situation requires, being responsible of one's work, taking initiative, being able to question one's own decisions as well as others, critical thinking, and asking for help.

4.5 Necessary societal data skills

Based on the societal interviews, we summarize societal data skills, why they were deemed important, and how could they be taught in Table 3. It is worth noting that some of the interviewees pointed out that teaching societal data skills should take place much earlier than in higher education.

In addition to our six preconceived categories, the interviewees also added *taking breaks from consuming information* and *understanding data life cycle* as necessary societal data skills. The former skill was seen related to managing information overload, and the latter to understanding how digital data is difficult to erase, how large corporations collect and analyze personal data, and how information is power on a global level.

5 Discussion

5.1 Comparison with prior findings

As we used the same framework of 24 data system skills as a previous study [25], comparing the findings is relatively effortless. From an international perspective (i.e., results reported previously), SQL, data modeling, cloud computing, and data pipelines were the most important skills in industry, and these findings are in line with the findings reported in our study from the Finnish industry. The importance of object-oriented data models, DBMS components, and relational algebra was deemed low in both Finnish and international industries. There were also some discrepancies. Namely, data warehousing and data visualization were highly sought skills in the Finnish industry, but not on the international level. In contrast, NoSQL DBMSs as well as functions and stored procedures were valued on the international level, but not on the Finnish level. The lack for the need for NoSQL DBMSs in Finland might be explained by systems developed for smaller number of end-users.

The other necessary skills reported by our study participants included specific technologies (e.g., vector databases and Parquet), understanding the landscape around data systems (e.g., software development frameworks and continuous development), and various soft skills (e.g., working with stakeholder and collaboration). The need for such soft skills has been reported in a large number of studies in prior literature [cf. e.g., 7, 26]. Similarly, understanding the holistic system development landscape has been broadly studied, especially through capstone courses [cf. e.g., 9].

As for the societal data skills, we are not aware of studies with similar research goals, i.e., providing *empirical data* on why such societal data skills are important or how to teach them in practice. However, the findings reported in Table 3 are in line with prior scientific literature. For example, the skill of identifying mis- and disinformation is important because such false information is used to influence world politics [10, 24]. That being said, the findings in Table 3 should be considered as a condensed source of both motivation for teaching societal skills, as well as concrete examples

Table 3: Societal data skills, why the skills are deemed necessary, and how they can be taught

Skill	Why the skill is deemed necessary	How to teach
Critical evaluation of data quality	Data cannot be reliably used otherwise; Data can be used to mislead; The trustworthiness of data has declined significantly; Emotional responses skew the interpretations	Identify reliable data sources; Teach critical thinking; **Exercise**: try to find information supporting a view you disagree with
Identifying mis- and disinformation	Data can be used to influence attitudes, usually in a negative way; No one is safe from misinformation or disinformation; Such information is effortless to fabricate; Factual information can be used in unsuitable contexts; Information overload leads to people making mistakes	What is the difference between misinformation and disinformation; Why they are used; **Exercise**: case-study, who influenced who, why, and with what kind of information
Interpreting and comparing data	Interpretation gives data a meaning; Data are used in decision making	Teaching scientific research methods; Teaching through examples how personal bias affects judgment; Evaluating what are the pieces of empirical evidence behind claims; **Exercise**: analysis of competing hypotheses
Data-driven decision-making	Decisions may have significant implications; Decisions are always based on some data; Higher education often prepares for decision-making roles	**Exercise**: make a decision based on a dataset
Ethical use of data	Data needs to used objectively, even when out of one's comfort zone; Ethics protect both the participants and the researchers, i.e., data objects and data handlers; It is a common premise to strive for ethical conduct	How to store and share data ethically; Copyright; Data use in commercial purposes; **Exercise**: managing data that one finds controversial personally
Data privacy and sharing	All digital content is used for profiling, which is then used in influencing; Targeted influencing; Exploiting cognitive vulnerabilities; Important to understand why personal information is coveted; Understanding what you want to share about yourself and to what ends those pieces of data can be used; Everyone has the right to privacy and a duty to respect the privacy of others	Describing how some actors try to silence other actors; Describing known influence operations; **Exercise**: What to do when one is targeted by a hate campaign

from experts on how to teach these skills, rather than findings that support or contest prior findings.

5.2 Practical implications

Our findings are applicable in developing data systems curricula. Based on the results, we will start to redesign the data systems education in the university of the authors to more closely account for the needs of industry, as well as the modern information landscape.

Specifically, we will *(i)* redesign a data systems module of 25 ECTS (1 ECTS equals roughly 27 hours of student work) based on the necessity of skills reported in Fig. 1, *(ii)* take into account the necessity of different skills: a skill with low necessity may be taught on a more general level, while critical skills should receive in-depth teaching, *(iii)* consider that while some more theoretical skills were not valued in the industry, the role of the university is to provide theoretical foundations in addition to practical skills, *(iv)* collaborate closely with other departments and faculties to incorporate additional and societal skills (e.g., both *data governance* and *identifying disinformation* from the faculty of management and business) into our data systems module.

5.3 Limitations and threats to validity

Sample representativeness: Our sample of 117 survey respondents and 34 interviewees may not fully represent the breadth of the Finnish data systems industry. In addition, the number of companies represented in the dataset remains unknown, as this information was not collected for privacy reasons. As such, certain skills among the 24 examined could be disproportionately emphasized, while others may be underrepresented. Nonetheless, when compared to prior international work (e.g., with 30 survey respondents [25]), our study draws on a relatively large dataset. Given the overall size of the eligible population in Finland, we believe the findings provide a reasonably representative picture, though they should be interpreted with these limitations in mind.

Qualitative analysis: A potential pitfall in the conventional content analysis is misinterpretation of the free form text answers. We tried to mitigate misinterpretations by independent analyses.

Framework limitations: Although current, the list of 24 skills [25] used in the survey might omit some emerging skills. We tried to mitigate this by a free form text field to voice such needs.

Generalizability: The levels of necessity of the 24 skills may not generalize to other countries, or on a broader, international scale. In fact, the comparison in Section 5.1 shows that some of the results are not in line with a previous study [25]. As the goal of our study was to *particularize* the framework [25] to the context of Finland, generalizability was not a primary focus. However, the research design can be replicated in other regions.

6 Conclusion

Data systems education and industry needs may be misaligned for various reasons. In this study, we analyzed the importance of different technical and societal data systems skills in the Finnish data systems industry and higher education. The results show that SQL and data modeling remain valued skills in the industry. We also show *why* and *how* to teach societal data skills. These results are applicable in developing higher education curricula.

Acknowledgments

This work has been funded by City of Tampere, Finland. We thank all study participants for their valuable contributions.

References

[1] Roberto O Andrade, Walter Fuertes, María Cazares, Iván Ortiz-Garcés, and Gustavo Navas. 2022. An exploratory study of cognitive sciences applied to cybersecurity. *Electronics* 11, 11 (2022), 1692.

[2] Daniel Bar-Tal and Dan Jacobson. 1998. A psychological perspective on security. *Applied psychology* 47, 1 (1998), 59–71.

[3] Francine Berman, Rob Rutenbar, Brent Hailpern, Henrik Christensen, Susan Davidson, Deborah Estrin, Michael Franklin, Margaret Martonosi, Padma Raghavan, Victoria Stodden, et al. 2018. Realizing the potential of data science. *Commun. ACM* 61, 4 (2018), 67–72.

[4] Katy Börner, Adam Maltese, Russell Nelson Balliet, and Joe Heimlich. 2016. Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization* 15, 3 (2016), 198–213.

[5] Julia Brasse, Maximilian Förster, Philipp Hühn, Julia Klier, Mathias Klier, and Lars Moestue. 2024. Preparing for the future of work: a novel data-driven approach for the identification of future skills. *Journal of Business Economics* 94, 3 (2024), 467–500.

[6] Stephen Cass. 2022. SQL Should Be Your Second Language. *IEEE Spectrum* 59, 10 (2022), 20–21. doi:10.1109/MSPEC.2022.9915547

[7] Iain J Cruickshank, Nathaniel D Bastian, Jean R.S. Blair, Christa M Chewar, and Edward Sobiesk. 2024. Seeing the Whole Elephant - A Comprehensive Framework for Data Education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. ACM, Portland OR USA, 248–254. doi:10.1145/3626252.3630922

[8] Tianxi Dong and Jason Triche. 2020. A longitudinal analysis of job skills for entry-level data analysts. *Journal of Information Systems Education* 31, 4 (2020), 312.

[9] Alan J Dutson, Robert H Todd, Spencer P Magleby, and Carl D Sorensen. 1997. A review of literature on teaching engineering design through project-oriented capstone courses. *Journal of engineering education* 86, 1 (1997), 17–28.

[10] Jacqueline Evans. 2022. War in the Age of TikTok. *Russian Analytical Digest* 280 (2022), 17–19.

[11] ACM Data Science Task Force. 2021. *Computing competencies for undergraduate data science curricula*. Association for Computing Machinery, New York, NY, USA. ISBN: 9781450390606.

[12] The Joint ACM/AIS IS2020 Task Force. 2020. A Competency Model for Undergraduate Programs in Information Systems.

[13] Adrian Gardiner, Cheryl Aasheim, Paige Rutner, and Susan Williams. 2018. Skill requirements in big data: A content analysis of job advertisements. *Journal of Computer Information Systems* 58, 4 (2018), 374–384.

[14] Bahareh Ghodoosi, Tracey West, Qinyi Li, Geraldine Torrisi-Steele, and Sharmistha Dey. 2023. A systematic literature review of data literacy education. *Journal of Business & Finance Librarianship* 28, 2 (2023), 112–127. doi:10.1080/08963568.2023.2171552 arXiv:https://doi.org/10.1080/08963568.2023.2171552

[15] Hilkka Grahn and Toni Taipalus. 2025. Defining comprehensive cognitive security in the digital era: Literature review and concept analysis. *Journal of Information Warfare* 2 (2025).

[16] Bin Guo, Yasan Ding, Yueheng Sun, Shuai Ma, Ke Li, and Zhiwen Yu. 2021. The mass, fake news, and cognition security. *Frontiers of Computer Science* 15, 3 (2021), 153806.

[17] Fatih Gurcan and Nergiz Ercil Cagiltay. 2019. Big data software engineering: Analysis of knowledge domains and skill sets using LDA-based topic modeling. *IEEE access* 7 (2019), 82541–82552.

[18] Hsiu-Fang Hsieh and Sarah E Shannon. 2005. Three approaches to qualitative content analysis. *Qualitative health research* 15, 9 (2005), 1277–1288.

[19] Tzu-Chieh Hung and Tzu-Wei Hung. 2022. How China's cognitive warfare works: a frontline perspective of Taiwan's anti-disinformation wars. *Journal of Global Security Studies* 7, 4 (2022), ogac016.

[20] Muhammad Ishaq, Adnan Abid, Muhammad Shoaib Farooq, Muhammad Faraz Manzoor, Uzma Farooq, Kamran Abid, and Mamoun Abu Helou. 2023. Advances in database systems education: Methods, tools, curricula, and way forward. *Education and information technologies* 28, 3 (2023), 2681–2725.

[21] Janne Kauttonen, Umair Ali Khan, Lili Aunimo, Antti Nyqvist, and Aarne Klemetti. 2024. Topic mining for theses and job ads in ICT sector: can higher education institutes respond to job market demands?. In *Frontiers in Education*, Vol. 9. Frontiers Media SA, 1322774.

[22] Melissa K Kjelvik and Elizabeth H Schultheis. 2019. Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy. *CBE—Life Sciences Education* 18, 2 (2019), es2.

[23] Amruth N. Kumar and Rajendra K. Raj. 2024. Computer Science Curricula 2023 (CS2023): The Final Report. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 2* (Portland, OR, USA) *(SIGCSE 2024)*. Association for Computing Machinery, New York, NY, USA, 1867–1868. doi:10.1145/3626253.3633405

[24] Stephen McCombie, Allon J Uhlmann, and Sarah Morrison. 2020. The US 2016 presidential election & Russia's troll farms. *Intelligence and National Security* 35, 1 (2020), 95–114.

[25] Daphne Miedema, Toni Taipalus, Vangel V. Ajanovski, Abdussalam Alawini, Martin Goodfellow, Michael Liut, Svetlana Peltsverger, and Tiffany Young. 2025. Data Systems Education: Curriculum Recommendations, Course Syllabi, and Industry Needs. In *2024 Working Group Reports on Innovation and Technology in Computer Science Education* (Milan, Italy) *(ITiCSE 2024)*. Association for Computing Machinery, New York, NY, USA, 95–123. doi:10.1145/3689187.3709609

[26] Farhad Sadik Mohammed and Fezile Ozdamli. 2024. A systematic literature review of soft skills in information technology education. *Behavioral Sciences* 14, 10 (2024), 894.

[27] Guido Ongena. 2023. Data literacy for improving governmental performance: A competence-based approach and multidimensional operationalization. *Digital Business* 3, 1 (2023), 100050.

[28] Maria-Magdalena Popescu. 2021. Disinformation deconstructed-cognition security and digital control. *Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences and Law* 14, 1 (2021), 123–134.

[29] Robert M Ross, David G Rand, and Gordon Pennycook. 2021. Beyond "fake news": Analytic thinking and the detection of false and hyperpartisan news headlines. *Judgment and Decision making* 16, 2 (2021), 484–504.

[30] Toni Taipalus and Ville Seppänen. 2020. SQL education: A systematic mapping study and future research agenda. *ACM Transactions on Computing Education (TOCE)* 20, 3 (2020), 1–33. doi:10.1145/3398377

[31] Nirnaya Tripathi. 2025. NoSQL database education: A review of models, tools and teaching methods. *Journal of Systems and Software* (2025), 112391.

[32] Amit Verma, Kirill M Yurov, Peggy L Lane, and Yuliya V Yurova. 2019. An investigation of skill requirements for business and data analytics positions: A content analysis of job advertisements. *Journal of Education for Business* 94, 4 (2019), 243–250.