Integrating data science education and responsible AI in teacher training: A pilot study within STEAM-based graduate learning

Maria Meletiou-Mavrotheris[1], Yianna Danidou[1], Efi Paparistodemou[1], Alexandros Kofteros[1] and Dionysia Bakogianni[2]
[1]European University, Cyprus
[2]Karlstad University, Sweden
m.mavrotheris@euc.ac.cy

This pilot study explores the integration of data science and responsible AI within a STEAM-based graduate course for education majors, developed as part of the EU-funded DataSETUP project. Focusing on a single module—Responsible AI & Data Science: Ethics, Society, and Citizenship—the research examines a group-based activity in which student teachers collaboratively built an image classification model using Google's Teachable Machine. Participants engaged with core elements of the machine learning pipeline and reflected on issues such as data quality, algorithmic bias, and educational relevance. To analyze student learning, a five-dimension analytical framework was developed, synthesizing concepts from data science, statistics education, and critical data literacy. The framework captures both technical engagement and student teachers' evolving understanding of the social and ethical implications of data-driven AI decision-making. Findings indicate that focused, hands-on experiences within a targeted module can meaningfully enhance student teachers' data science literacy and their critical awareness of responsible AI use in education.

INTRODUCTION

The exponential growth of data—fueled by the widespread use of social media, digital technologies, and artificial intelligence (AI)—has profoundly reshaped how societies produce knowledge, make decisions, and respond to complex challenges. Data science, defined as "the study of extracting value from data" (Wing, 2019), now plays a pivotal role in fields as diverse as climatology, healthcare, education, and public policy (Biehler et al., 2018; Tanaka et al., 2022; Engel, 2017). While open and multisource access to data is a foundational principle of democratic societies, the inability to critically and ethically navigate this information landscape can lead to misinformation, manipulation, and disengagement from informed civic participation (Bobrowicz et al., 2022).

Despite the growing societal relevance of data science, research shows that many secondary school graduates struggle to interpret data, assess the credibility of sources, and detect bias in digital content (Konold et al., 2015; Wineburg et al., 2016; Zucker et al., 2020). These challenges underscore the urgent need to foster students' data science literacy—not only as a technical competence but also as a civic and ethical one (Gould, 2017; 2021). This need becomes even more pressing in light of AI's increasing role in decision-making. Educators must be prepared to teach students how data underpins Machine Learning (ML), how algorithmic decisions are made, and how to critically evaluate the fairness, transparency, and social implications of AI systems (Ridgway, 2016; Ridgway et al., 2022).

Despite growing recognition of its importance, structured approaches to integrating data science into school curricula are still largely absent—particularly within initial teacher education programs (Dunlap & Piro, 2016; Mandinach & Gummer, 2016; Henderson & Corry, 2021). Addressing this gap, the EU-funded Erasmus+ project DataSETUP (November 2023–October 2026) aims to embed data science education into university-level teacher preparation programs. Implemented by a multinational consortium of universities in Germany, Cyprus, Greece, Ireland, and Turkey, the project aligns with European priorities on digital transformation, civic engagement, and inclusive education (Maas et al., 2022; Makar, Fry & English, 2023).

Adopting a design-based research (DBR) methodology (Cobb et al., 2003), the DataSETUP project develops, tests, and refines educational modules, digital tools, and assessment frameworks that promote data science as both a technical and civic competency. Emphasizing not only data analysis and visualization but also the ethical, societal, and civic dimensions of data and AI, the project aims to prepare student teachers to engage learners in critical, justice-oriented STEAM education. This paper presents findings from the first pilot implementation of one such module—*Responsible AI & Data Science: Ethics, Society, and Citizenship*—delivered as part of the graduate course *AI in STEAM*

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*.

Education at European University Cyprus. Designed to introduce student teachers to foundational data-related concepts in AI and ML, the module explores the role of data in algorithmic decision-making and addresses key concerns related to bias, fairness, and the broader societal implications of data-driven AI systems. The study presented here investigates the extent to which participation in the module enhances student teachers' data science literacy and their preparedness to teach AI- and data-related topics in educational contexts.

METHODOLOGY

Participants and Context

AI in STEAM Education is an elective course of the online MA Program Technologies of Learning & Communication and STEAM Education at European University Cyprus. Delivered via the Blackboard platform, the course comprises 13 modules, including two focused on AI/ML-related data science: the *Responsible AI & Data Science: Ethics, Society, and Citizenship* module examined in this study, and a subsequent module titled *Integrating AI/ML-Related Data Science in STEAM Education to Foster Responsible and Active Citizenship*. The course provides student teachers with study materials and tools to support both synchronous and asynchronous communication and collaboration.

Offered for the first time in Spring 2025, *AI in STEAM Education* enrolled 36 student teachers, all of whom participated in this study. The vast majority of participants were female (n=32, 88.9%), and two-thirds were over the age of 30, with nearly half falling within the 31–40 age range. Most had at least 1–2 years of teaching experience, and approximately one-third had taught for more than five years. Regarding their teaching levels, about half were teaching or preparing to teach at the pre-primary or primary level, while the other half were involved in secondary education. The participants' academic backgrounds were diverse, with the most common areas of specialization including Language and Literature, Computer Science and Engineering, Early Childhood Education, Primary Education, and Special and Inclusive Education.

The *Responsible AI & Data Science: Ethics, Society, and Citizenship* module was designed to equip the student teachers enrolled in the course with the knowledge, skills, and dispositions needed to critically engage with AI and data science in education. Targeting a diverse cohort of participants, the module invites student teachers to explore how data and algorithms shape automated decision-making in ways that raise significant pedagogical and societal questions.

The module combines instructor-led input with experiential, inquiry-based activities. It begins by exploring the presence of AI in daily life and introducing key concepts in ML and data-driven decision-making. Activities such as *ChatGPT in Action* prompt student teachers to assess AI-generated responses for accuracy, bias, and transparency, encouraging critical reflection on how large language models (LLMs) like ChatGPT generate output based on probabilistic patterns.

Instruction incorporates the statistical investigation cycle (PPDAC) and connects it to AI problem-solving processes (Wild & Pfannkuch, 1999), while also highlighting how data science extends traditional statistical thinking by working with complex, often pre-existing datasets and by emphasizing modeling, iteration, and ethical implementation. Through this lens, student teachers are introduced to issues such as algorithmic bias, data ethics, and the societal impacts of AI.

Hands-on, inquiry-based learning forms the pedagogical core of the module. Student teachers use tools such as Google's *Teachable Machine* and *Machine Learning for Kids* to collaboratively design, train, and evaluate image classification models—developing practical skills in data preparation (collection, labeling, cleaning) and critically reflecting on the ethical implications of algorithmic decision-making. Through activities like building gesture-recognition games or engaging with Code.org's *AI for Oceans*, student teachers explore how bias, transparency, and fairness affect the accuracy and reliability of AI systems, and how these systems can either reinforce or challenge existing inequities. The module seamlessly combines conceptual learning with hands-on design to cultivate critical data science literacy and ethical awareness within a STEAM education context.

Data Collection and Analysis

A mixed-methods approach informed the data collection process. Multiple forms of both qualitative and quantitative data were gathered, including pre- and post-surveys measuring participants' attitudes, perceptions, and self-efficacy related to data science education; classroom observations capturing engagement and instructional flow; teacher educator reflection sheets; student workbooks documenting learning experiences; and online discussions supporting collaborative reflection. While these sources offered a broad perspective on the module's impact, this paper focuses specifically on a group activity conducted within the module. In this task, student teachers collaborated in eight groups of 4–5 to design and develop a basic image classification model using Google's *Teachable Machine* platform. They were tasked with designing a model by collecting, labeling, and training image datasets. Student teachers engaged with core elements of the ML pipeline, such as data collection and cleaning, class balancing, training-validation testing, and performance evaluation. They also completed structured workbooks prompting them to reflect on data quality, the influence of biased inputs, ethical concerns in AI decision-making, and the pedagogical relevance of ML in classroom settings. The activity fostered both technical skill-building and ethical reflection, offering a rich site for investigating student learning.

To analyze student responses and workbooks, we developed the analytical framework and coding schema summarized in Table 1.

Table 1. Refined Analysis Framework with Extensions

Framework Dimension	**Analytical Questions**	**Indicative Codes (Coding Categories)**
1. Asking Questions with Data	• What question is being addressed? • Is the question socially, contextually, or pedagogically grounded? • Does the question have educational value for students?	• Problem formulation • Purpose and relevance • Connection to students' lives or curriculum
2. Collecting, Cleaning, and Manipulating Data	• How are data selected, labeled, or framed? • Is there a reflection on the origin, structure, or implicit bias in the data? • Do student teachers engage with the 'messiness' of real data?	• Data diversity and framing • Representational bias awareness • Strategies to reduce bias
3. Modelling and Interpreting	• What methods are used and why? • How are results contextualized and explained? • What types of reasoning (statistical, algorithmic, interpretive) are shown?	• Performance evaluation • Reflective analysis • Purpose-driven modeling choices
4. Critiquing Data-Based Claims	• Are claims supported and critiqued? • Are issues of fairness, validity, and interpretation addressed? • Is there reflection on the implications of automation?	• Bias detection • Socio-ethical discussion • Openness to limitations
5. Data Epistemology	• Where do data come from and how were they produced? • What assumptions do the data carry? • What knowledge is constructed through data?	• Data origin and authorship • Visibility/invisibility of perspectives • Epistemic stance toward data

The five-dimension analytical framework synthesizes themes from data science education, statistics education, and critical data literacy. The first four dimensions—(1) Asking Questions with Data, (2) Collecting, Cleaning, and Manipulating Data, (3) Modeling and Interpreting, and (4) Critiquing Data-Based Claims—draw on inquiry-based learning goals outlined in the GAISE II framework (Bargagliotti et al., 2020), the Data Science Course Framework developed by the Charles A. Dana Center (2022), and emerging research in computational thinking in education (e.g., Lee et al., 2022; Kazak et al., 2025). These dimensions emphasize student teachers' ability to formulate meaningful, contextually relevant questions; work with messy, real-world data; interpret outcomes; and critically evaluate claims in light of fairness, bias, and ethical considerations. Together, they offer a structured lens for examining how student teachers engage with core data practices and begin to envision their pedagogical application.

The fifth dimension—Data Epistemology—was introduced inductively based on student teachers' reflective responses and is grounded in critical data studies. This dimension captures awareness of how data are socially constructed, whose perspectives are represented or excluded, and what assumptions underpin data generation and representation. It is informed by scholarship that highlights the ideological and political nature of data practices, particularly in educational contexts (Philip et al., 2016; D'Ignazio & Klein, 2020), underscoring the importance of interrogating the power structures and values embedded in the production and use of data.

RESULTS

Engagement Across Analytical Framework Dimensions

Analysis of group workbooks and reflections revealed varying levels of engagement across the five dimensions of the analytical framework. While most groups demonstrated technical competence with core machine learning processes, the depth of critical reflection—especially regarding data epistemology and ethical critique—was uneven:

1. Asking Questions with Data: Across several projects, student teachers formulated technically clear classification tasks (e.g., distinguishing between fruits, emotions, or vehicles). However, only a few groups (notably Groups 1, 7, and 8) explicitly connected their inquiry to pedagogical or societal relevance. This highlights a tendency to treat data science as a technical task, rather than as a tool for exploring meaningful, context-rich questions. Moreover, several groups defined classification problems (e.g., 'orange' vs. 'not orange', types of vehicles), but only a few embedded these tasks in meaningful educational contexts. Here is an example from Group 7: *"...the activity was implemented for educational purposes for preschool children... categories included apple, banana, orange, grapes, and strawberry."*

2. Collecting, Cleaning, and Manipulating Data: Most groups demonstrated an awareness of the importance of data diversity and balance, often referring to variations in lighting, angles, and backgrounds. Several discussed common challenges in dataset construction, such as bias in image sources and label imbalance. One group (Group 4), however, went further by explicitly acknowledging issues of cultural representation. In their reflection, they noted: *"Using only professional food photos could misrepresent everyday eating habits... cultural identity is shaped by what is and isn't labelled healthy."* This comment reflects a critical awareness of how data choices are not neutral; rather, they participate in shaping how students, teachers, and even ML models perceive and categorize the world. It highlights how the construction of datasets can reinforce particular social assumptions and potentially marginalize alternative perspectives. Despite this, few groups engaged deeply with the implications of data framing or labelling in terms of epistemology or power.

In Group 6's essay, for instance, the focus remained more technical: *"We included images of different lighting, background, and angle... however, misclassification occurred with visually similar fruits like grapefruit and peach."*

3. Modelling and Interpreting: Most groups successfully built and evaluated classification models using *Teachable Machine*. They generally conducted performance testing and reflected on unexpected outcomes, with some offering thoughtful analysis of specific errors.

In several cases, student teachers did more than report accuracy; they attempted to interpret why certain errors occurred. Group 3, for instance, presented a detailed performance table (see Table 2) that revealed the model's difficulty in distinguishing between visually similar fruits. Half peaches and grapefruits were frequently misclassified as oranges—an issue they linked to shared visual characteristics such as color and shape.

Table 2. System Evaluation: Classification Accuracy by Object (%)

Object	**Classified as "Orange"**	**Classified as "Not Orange"**
Half Peach	88%	12%
Whole Peach	19%	81%
Grape	3%	97%
Half Pear	2%	98%
Grapefruit	100%	0%

This kind of interpretation was also evident in Group 6's robustness testing. They described how classification accuracy dropped significantly under multicolored backgrounds and low lighting conditions, reporting a 46% error rate in such cases. Their observation—*"The model showed a 46% error rate in images of oranges with multicolored background and low lighting"*—demonstrates an awareness that model evaluation must go beyond ideal conditions.

In a few cases, modeling choices were clearly shaped by educational intent. Group 1, for example, developed a model to classify facial expressions of joy and anger—emotions they selected for their clear visual contrast and relevance to classroom interactions. While the task introduced them to key ML concepts, the group also reflected on how such a model could support emotional awareness in students and help teachers respond more effectively to learners' emotional cues. This suggests an emerging capacity to align data science activity with pedagogical goals.

Still, some groups remained focused primarily on surface-level accuracy metrics, without fully interpreting the implications of the model's behavior in real-world or classroom contexts.

4. Critiquing Data-Based Claims: Across the projects, student teachers demonstrated growing awareness of fairness, bias, and the socio-ethical implications of algorithmic decision-making. Several groups went beyond technical evaluation to reflect on the social impact and ethical responsibility of their models. Groups 2, 4, and 8 in particular offered thoughtful critiques of model limitations, the risks of misclassification, and the consequences of applying AI in educational or real-world contexts.

Group 4, for example, questioned how the use of only professional food photos might reinforce narrow cultural representations of what counts as 'healthy', implicitly pointing to how labeling choices may reflect social values rather than neutral truths. Group 8 addressed the risk that biased training data could reinforce stereotypes, emphasizing that *"it's important to ensure responsible and fair use of image classification models in education".* This comment reveals an awareness not only of bias in datasets, but of the need to approach modeling choices with social care and pedagogical purpose.

Some student teachers also acknowledged the limitations of their models in practice. Rather than presenting them as infallible, they openly discussed when and why the model failed, and what that might mean for its application in classrooms or broader decision-making contexts.

Still, in most projects, such critiques remained abstract or disconnected from learners' actual experiences, highlighting the need to further support student teachers in translating ethical awareness into meaningful classroom discussions and responsible teaching practice.

5. Data Epistemology: Only a few groups engaged with the deeper epistemological dimension of data science—that is, questioning what counts as data, who defines categories, and how these choices shape the knowledge produced by a model. Most groups treated data as neutral, objective inputs, without considering how datasets are socially constructed and embedded with assumptions.

A notable exception was Group 4, who reflected on how the use of curated, professional food images might distort real-world eating practices. Their comment — *"Using only professional food photos could misrepresent everyday eating habits... cultural identity is shaped by what is and isn't labeled healthy"*—demonstrates a critical stance toward data selection and labeling. Student teachers implicitly recognized that the origin and authorship of the data (images sourced from professional contexts rather than students' lived experiences) affect the kinds of realities that the model represents.

At the same time, Group 4 drew attention to the invisibility of certain perspectives, such as every day or culturally specific food practices, which are often excluded from standard datasets. This reflection shows that the student teachers are beginning to see data not as neutral facts, but as choices shaped by people, context, and culture—choices that should be questioned and interpreted critically.

Such moments underscore the importance of helping student teachers see data not only as information to be analyzed, but also as cultural artifacts that reflect—and sometimes obscure—diverse ways of knowing.

Connections to Teaching Practice

Although the assignment did not explicitly require pedagogical integration, several groups made spontaneous connections between their data science tasks and classroom practice. These insights illustrate how data science activities can prompt student teachers to consider instructional use, even in the absence of direct prompting.

Group 1 proposed using emotion-recognition models to support socio-emotional learning. Group 8 envisioned using transportation classification as a springboard for interdisciplinary STEAM lessons. Group 7, framed their fruit classification model as suitable for preschool education, highlighting early learning potential around categorization and vocabulary.

These examples reflect an emerging pedagogical lens: student teachers began to see themselves not just as builders of AI models, but as educators exploring how these tools could support learners. However, such connections were limited and varied in depth, suggesting a need for more structured opportunities within teacher education to link technical activities to instructional design.

The emergence of these unprompted pedagogical reflections reveals the transformative potential of integrating responsible AI and data science within teacher preparation. They suggest that hands-on, ethically oriented tasks can do more than build skills—they can help student teachers envision how to *teach with* and *about* AI/ML data science in meaningful, context-sensitive ways.

DISCUSSION AND CONCLUSION

This study highlights the potential of a targeted, hands-on module within a STEAM-based graduate course to foster technical proficiency and ethical awareness in data science education. Drawing on a structured five-dimension framework, we found that student teachers engaged meaningfully with the technical and ethical dimensions of AI and data science—particularly in formulating data-driven questions, designing image classification models, and reflecting on fairness, bias, and model performance.

Student teachers displayed increasing sophistication in the first three analytical dimensions: formulating meaningful, contextually grounded questions (Asking Questions with Data); engaging with data preparation tasks such as collecting, cleaning, and organizing datasets (Collecting, Cleaning, and Manipulating Data); and building simple models while interpreting their performance and limitations (Modelling and Interpreting). Many groups also demonstrated an emerging capacity in the fourth dimension—Critiquing Data-Based Claims—by recognizing algorithmic bias and considering the social consequences of automated decisions. However, deeper engagement with the sociopolitical nature of data—captured in the fifth dimension of Data Epistemology—was limited to a small number of groups. This suggests that while student teachers were open to critical reflection, they may lack conceptual tools or contexts rich enough to prompt deeper interrogation of data's constructed nature. Notably, this limitation may also stem from the simplicity of the ML tasks involved. The activity's use of *Teachable Machine* to create basic image classifiers likely constrained student teachers' opportunities to encounter the ambiguity, partiality, or contested nature of real-world data.

These findings highlight the importance of designing data science education experiences that go beyond surface-level technical engagement. While technically accessible platforms like *Teachable Machine* are valuable entry points, more complex, socially grounded data tasks may be necessary to provoke sustained reflection on whose data is being used, how categories are defined, and what social values or exclusions are embedded in data collection and labeling practices.

An additional insight emerged through student teachers' spontaneous connections to teaching practice. Although pedagogical integration was not explicitly required in the assignment, several groups imagined how their data science activities might be adapted for educational contexts. These connections suggest that hands-on, critically framed tasks not only support student teachers' learning, but also encourage them to envision the classroom relevance of AI and data literacy. This pedagogical dimension—initially considered as a sixth analytical category but later separated from the core framework—offers a promising area for further research. Understanding how student teachers begin to translate technically focused tasks into instructional designs could provide a critical link between theory, practice, and teacher professional identity.

To this end, the follow-up module in the *AI in STEAM Education* course—*Integrating AI/ML-Related Data Science in STEAM Education to Foster Responsible and Active Citizenship*—is explicitly designed to deepen these connections. It guides student teachers in designing classroom-ready STEAM activities and full educational scenarios, aligning data science with curriculum goals, and engaging learners in justice-oriented inquiry using real-world datasets. By bridging ethical AI with practical teaching strategies, the second module builds on the conceptual foundation established here and deepens student teachers' capacity to teach with and about data in socially meaningful ways.

Overall, this study underscores that effective data science education for student teachers must integrate three interdependent strands: technical skills, critical reflection, and pedagogical translation. Supporting student teachers in moving fluidly between these domains can enable them not only to understand data science, but to teach it in ways that are responsive to learners' lives, social contexts, and the ethical demands of our digital age.

REFERENCES

Bargagliotti, A. E., Franklin, C., Arnold, P., Gould, R., & Miller, D. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education*. American Statistical Association. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf.

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 139–192). Springer. https://doi.org/10.1007/978-3-319-66195-7_5.

Bobrowicz, K., Han, A., Hausen, J., & Greiff, S. (2022). Aiding reflective navigation in a dynamic information landscape: A challenge for educational psychology. *Frontiers in Psychology, 13*, 881539. https://doi.org/10.3389/fpsyg.2022.881539.

Charles A. Dana Center at The University of Texas at Austin. (2024). *Data science course framework*. https://www.utdanacenter.org/frameworks.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13. https://doi.org/10.3102/0013189X032001009.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press. https://doi.org/10.7551/mitpress/11805.001.0001.

Dunlap, K., & Piro, J. S. (2016). Diving into data: Developing the capacity for data literacy in teacher education. *Cogent Education, 3*(1), 1132526. https://doi.org/10.1080/2331186X.2015.1132526.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal, 16*(1), 44–49. https://doi.org/10.52041/serj.v16i1.213.

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal, 16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209.

Gould, R. (2021). Toward data-scientific thinking. *Teaching Statistics, 43*(S1). https://doi.org/10.1111/test.12267.

Henderson, J., & Corry, M. (2021). Data literacy training and use for educational professionals. *Journal of Research in Innovative Teaching & Learning, 14*(2), 232–244. https://doi.org/10.1108/JRIT-11-2019-0074.

Kazak, S., Leavy, A., & Podworny, S. (2025). Framework for developing data science education modules for teacher education in STEAM subjects [Poster presentation]. *14th Congress of the European Society for Research in Mathematics Education (CERME14)*, Bolzano, Italy.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics, 88*(3), 305–325. https://doi.org/10.1007/s10649-013-9529-8.

Lee, H., Mojica, G., Thrasher, E., & Baumgartner, P. (2022). Investigating data like a data scientist: Key practices and processes. *Statistics Education Research Journal, 21*(2). https://doi.org/10.52041/serj.v21i2.41.

Maass, K., Sorge, S., Romero-Ariza, M., Hesse, A., & Straser, O. (2022). Promoting active citizenship in mathematics and science teaching. *International Journal of Science and Mathematics Education, 20*, 727–746. https://doi.org/10.1007/s10763-021-10182-1.

Makar, K., Fry, K., & English, L. (2023). Primary students' learning about citizenship through data science. *ZDM Mathematics Education, 55*, 967–979. https://doi.org/10.1007/s11858-022-01450-7.

Mandinach, E. B., & Gummer, E. S. (2016). Every teacher should succeed with data literacy. *Phi Delta Kappan, 97*(8), 43–46. https://doi.org/10.1177/0031721716647018.

Philip, T. M., Olivares-Pasillas, M. C., & Rocha, J. (2016). Becoming racially literate about data and data-literate about race: Data visualizations in the classroom as a site of racial-ideological micro-

contestations. *Cognition and Instruction, 34*(4), 361–388. https://doi.org/10.1080/07370008.2016.1210418.
Ridgway, J. (2016). Implications of the data revolution for statistics education. *International Statistical Review, 84*(3), 528–549. https://doi.org/10.1111/insr.12110.
Ridgway, J., Campos, P., & Biehler, R. (2022). Data science, statistics, and civic statistics: Education for a fast changing world. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement: Teaching civic statistics to develop informed citizens* (pp. 563–580). Springer. https://doi.org/10.1007/978-3-031-20748-8_22.
Tanaka, T., Himeno, T., & Fueda, K. (2022). University's endeavor to promote human resources development for data science in Japan. *Japanese Journal of Statistics and Data Science, 5*, 747–755. https://doi.org/10.1007/s42081-022-00151-5.
Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x.
Wineburg, S., McGrew, S., Breakstone, J., & Ortega, T. (2016). *Evaluating information: The cornerstone of civic online reasoning*. Stanford Digital Repository. http://purl.stanford.edu/fv751yt5934.
Wing, J. M. (2019). The data life cycle. *Harvard Data Science Review, 1*(1). https://doi.org/10.1162/99608f92.e26845b4.
Zucker, A., Noyce, P., & McCullough, A. (2020). JUST SAY NO! Teaching students to resist scientific misinformation. *The Science Teacher, 87*(5), 24–29. https://doi.org/10.2505/4/tst20_087_05_24.