Empowering Teachers to Foster Data Science Literacy for Civic Engagement and Social Justice

DataScEd4CiEn

Maria Meletiou-Mavrotheris, Efi Paparistodemou, Ana Serradó-Bayés, Dionysia Bakogianni

6th International Conference "Literacies and Contemporary Society: From Skills to Practices"
24–25 February 2026
Nicosia, Cyprus

1

The Data Landscape and Its Challenges

DataScEd4CiEn

- The vast increase in data availability through digital media and AI provides valuable insights across all life sectors
- Open access to diverse information sources supports democratic values but requires navigation skills to prevent misinformation
- Research shows most secondary school graduates struggle with
  - Interpreting and assessing scientific and political data validity (Konold et al., 2015; Zucker et al., 2020; Kahne & Bowyer, 2017)
  - Detecting bias in online sources (Wineburg et al., 2016)
- **Implication:** There is a critical need to foster data science literacy to empower informed, civic-minded citizens.

2

Importance of data science education

DataScEd4CiEn

- **Definition:** Data science is "the study of extracting value from data" (Wing, 2019), essential for tackling important societal issues like climate change and public health
- Increased demand for a new generation of scientists who can anticipate user needs and develop optimal solutions to address business, academic and societal challenges (Seshaiyer & McNeely; 2022)
- Introduction to applied data science from an early age can help create a new generation of highly skilled data scientists (Irizarry; 2020)
- Data science literacy is becoming increasingly important for every profession, from technology, science, finance, journalism and politics to art and history (Mund, 2022)
- Engaging with data science through problems relevant to society prepares students to cope with their role as active, responsible, thinking members of society

Co-funded by the European Union

3

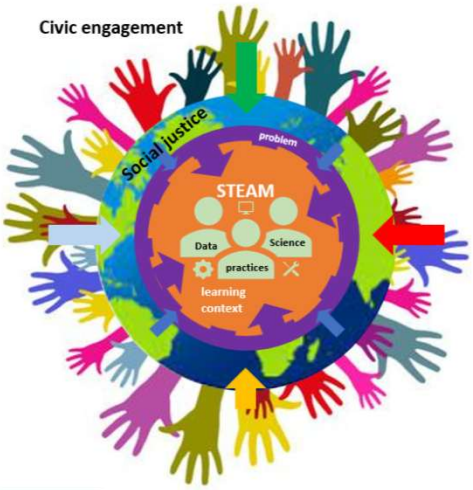
Current situation

DataScEd4CiEn

- Limited supportive material for data science in teacher education (Batanero et al. 2011)
- Limited opportunities for pre-service and in-service teacher training in data science education

Co-funded by the European Union

4



Project Overview - DataScEd4CiEn

- **Focus:** Data science education in STEAM for civic engagement and social justice from early years (ages 9-15)
- **Goals:** Establish core ideas for data science education, equip teachers with necessary skills and resources, develop educational materials, and conduct impact research

DataScEd4CiEn

5

Project Partners:

- European University Cyprus (EUC)
- The English School (ES) Cyprus
- National Kapodistrian University of Athens (UNKA) Greece
- Westfaelische Wilhelmsuniversitaet (WWU) Germany
- Mary Immaculate College University of Limerick (MIC) Ireland
- La Salle-Buen Consejo (LSBC) Spain

Researchers with a substantial experience and prior work in several related areas:

- STEAM
- Data Science Education
- Teacher education & professional development
- Early statistical and probabilistic thinking

Co-funded by the European Union

6

AIM of the project

DataScEd4CiEn

THE THEORETICAL FRAMEWORK to establish a set of core ideas for data science education for students aged 9-15 years

PROFESSIONAL DEVELOPMENT PROGRAM to equip STEAM teachers with the necessary knowledge, skills and resources to bring data science into the classroom

INSTRUCTIONAL MATERIALS to empower teachers to combine and implement STEAM scenarios to promote student social awareness through data science; and

RESEARCH to evaluate the impact on teachers' professional knowledge and on students' attitudes and learning.

Co-funded by the European Union

7

DataScEd4CiEn

Teacher PD model: Modular blended PD + Communities of Practice

Interactive workshops and case-based discussions

Hands-on exploration of tools (examples: **CODAP**, Gapminder, Teachable Machine, Machine Learning for Kids, GeoGebra, Excel, open data portals)

Co-design studio: teachers collaboratively build **data-rich civic STEAM scenarios**

Move from fundamentals → classroom-ready enactment

Co-funded by the European Union

8

DataScEd4CiEn

Stages of the project- *Design Based Research*

(Hußmann et al. 2013)

External Evaluator

- Development of theoretical framework
- Development of a research-action plan
- Design of tools and resources: professional development course, STEAM Scenarios, E-Platform

Design

Hypothesis

Piloting in partner/associated partner schools

Feedback/ Improvement

- Pilot testing of professional development course and classroom experimentation in partner schools
- Evaluation of materials
- Evaluation of methods
- Consistency check
- Development of research questions

Finalisation

- Large-scale implementation of professional development course
- Research and Evaluation of Materials
- Research and Evaluation of Methods
- Adaptation to national framework

Large scale Implementation

Dissemination

- National reports
- External reviewer report
- Final project report
- Publications in high impact journals and conferences
- E-Platform as OEA

Co-funded by the European Union

9

Co-funded by the European Union

Design-Based Research

Methodology and data sources

DBR Cycle 1 (2024–2025)

- Teacher PD → co-design of scenarios → classroom enactment → reflection & refinement
- Triangulated mixed methods across countries
- Student study in ICOTS-12 paper: 277 students (Cyprus, Greece, Spain), ages 11–16

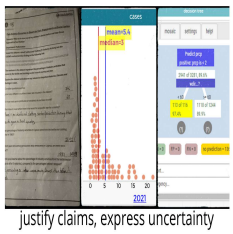
Instruments (WP6)

- Pre/post surveys (teachers & students)
- Scenario design artifacts, student outcomes
- Teacher diaries, classroom observations/video, focus groups

DataScEd4CiEn → **Data collection**

Primary basis of argumentation

Framing of tasks



justify claims, express uncertainty

Opportunities for subjective or open-ended reasoning

Source: Toulmin-based lens (SRTL-14) to study subjectivities & perspectives (Meletiou-Mavrotheris et al., 2025)

10

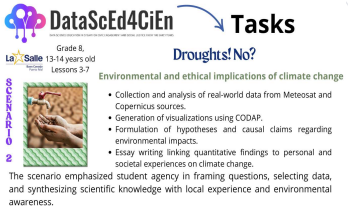
Classrooms

What it looked like in classrooms

Three snapshots of "data + meaning-making + action"

Fast Fashion (Cyprus)

Eurostat data + CODAP; student arguments about environmental/ethical "true cost".



Drought (Spain)

Local climate data + CODAP; hypotheses, causal claims, and design-thinking solutions.

Reducing Food Waste: Learning to Act with Data (Greece)

Food waste + CODAP with authentic data; dispositions: curiosity, ethical reasoning, empathy, civic agency

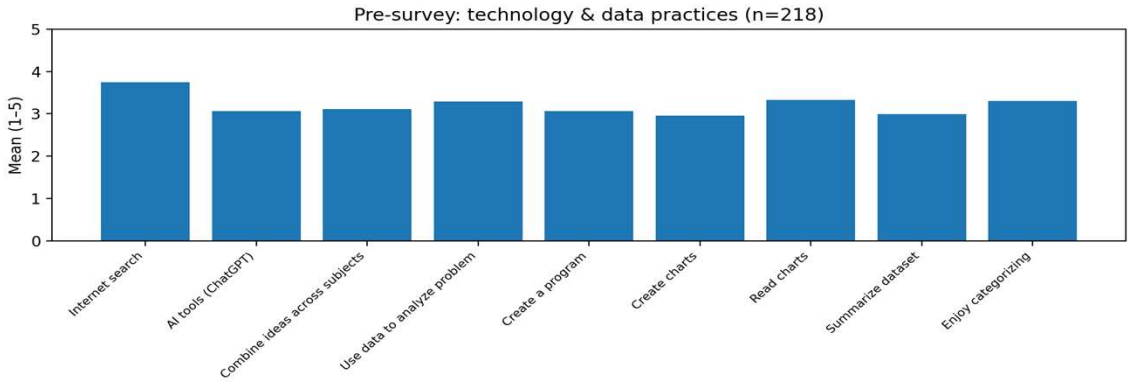
Co-funded by the European Union

11

Students

Where students started (pre-survey)

Self-reported practices before scenarios (n=218; 5-point scale)



Key gap: students reported higher confidence interpreting charts than producing displays / summarizing data.

Co-funded by the European Union

12

DataScEd4CiEn

Content analysis — How we analyzed the 8 scenario designs

Data source

- 8 STEAM Scenarios on data science for civic engagement and social justice created by teacher teams (Cycle 1)

Purpose

- Identify *which* data-science practices and *which* technologies teachers embedded in co-designed STEAM scenarios
- Make visible the **developmental progression** (ages 11–16) in both data practices and tool sophistication

Coding dimensions (technology + data practices)

- **Digital tools** used (e.g., CODAP, Excel/GeoGebra, Eurostat/World Bank, Global Forest Watch, AI tools)
- **Data moves**: filtering, grouping, summarizing, calculating new variables, joining datasets
- **Exploratory Data Analysis (EDA)**: descriptive statistics, visual exploration, patterns/outliers
- **Reasoning with data**: interpreting results in context; drawing conclusions; making judgments/decisions
- **Predictive modelling**: regression, decision trees / ML-based prediction
- **Big data ideas**: volume/variety/veracity/value; sampling as a response to scale/complexity
- **Programming/coding**: automation + algorithmic thinking (primarily block-based environments)

Output of analysis

- A cross-scenario map of **what practices appear where**, and **how complexity increases with age**

Co-funded by the European Union

13

Design

DataScEd4CiEn

Co-funded by the European Union

What teachers designed: a developmental trajectory

Content analysis of 8 teacher-designed scenarios (ICOTS-12 paper)

Target ages: 11–13	Foundational data work & early computational experiences Emphasis on interpreting/representing data, often with block-based tools (e.g., Scratch/ML for Kids)
Target ages: 13–15	More explicit data moves & multi-tool analysis workflows Increased use of filtering/grouping/calculating variables alongside EDA
Target ages: 15–16	Predictive modelling (regression, decision trees) Big-data focus framed through sampling and limitations of working with large/complex datasets

Tool ecosystem (what teachers relied on):

- CODAP as the "anchor" platform in most scenarios → continuity across countries/ages;
- Contextual data sources added as needed (e.g., Eurostat, World Bank, national repositories, Global Forest Watch);
- Other tools supported specific tasks (e.g., Excel/GeoGebra, occasional ChatGPT)

Across scenarios:

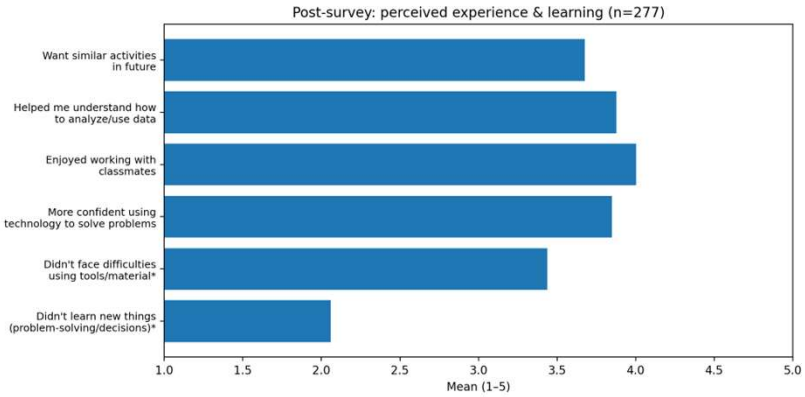
- EDA + reasoning with data were universal; modelling appeared only in oldest group; CODAP was the core analytic platform.

So what? Teachers' designs suggest a scalable pathway: *EDA → data moves → modelling, while keeping civic meaning-making central*

14

Findings

How students experienced the scenarios (post-survey)



- Enjoyed collaboration (high)
- Increased confidence using technology to solve problems (high)
- Perceived improved ability to analyze/use data (high)
- **Equity signal:** No meaningful age- or gender-related differences in perceived learning/engagement→ suggests inclusive scenario design can dampen typical equity gaps (in this context).

Co-funded by the European Union

15

Takeaways

Cross-country implementation patterns (what we learned)

Common strengths

- High student engagement when inquiry is anchored in real civic issues
- EDA and reasoning with data were central across scenarios
- Interdisciplinary collaboration enabled “data and meaning” (not data alone)

Common challenges

- Time constraints (PD + classroom schedules)
- Platform/tool usability + language accessibility
- Need for curated datasets, localized resources, mentoring

Co-funded by the European Union

16

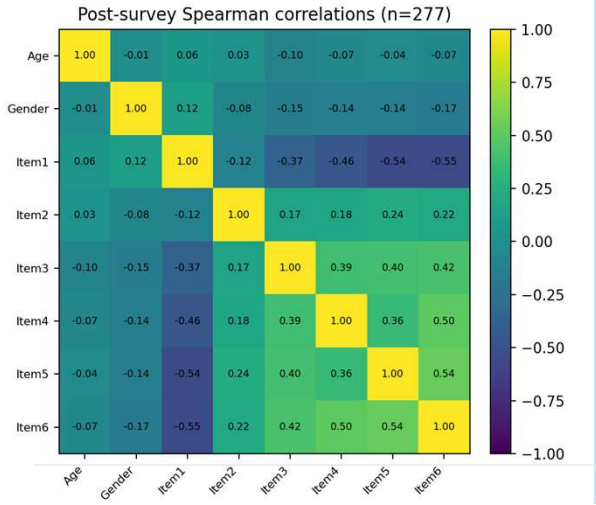
Co-funded by the European Union

DataScEd4CiEn

Key tension

A critical challenge: "technical" gains vs. epistemic recognition

Spearman correlations from post-survey (Serrado et al., 2026)



What the main variables mean

- **Age, Gender:** demographic variables
- **Item 3:** technology-related confidence
- **Item 5:** understanding / improvement in data analysis
- **Item 1 (reverse-coded for analysis):** originally negatively worded ("I did NOT gain new general problem-solving/decision-making skills")→ reverse-coded so higher Item 1 = more perceived general/transferable problem-solving & decision-making gains

Key results

- **Age:** all correlations trivial ($|\rho| \leq 0.104$) → younger/older students perceived the activity similarly
- **Gender:** very small correlations (e.g., Item 3: $\rho = -0.149$; Item 5: $\rho = -0.142$)
- **Strongest relationship:** Item 1 vs Item 5: $\rho = -0.540$ ($p < .001$)→ students with stronger perceived data-analysis gains were least likely to report general decision-making/problem-solving gains
- **Also notable:** Item 3 vs Item 5 positive ($\rho \approx 0.40$)→ tech confidence aligns with perceived data-analysis improvement

Interpretation

- Students can enjoy and learn data practices (analysis/modelling tools) without seeing them as a general, transferable epistemic tool for decisions

17

DataScEd4CiEn

Takeaways

What we learned

Three actionable takeaways

- Teacher PD works when it is collaborative, interdisciplinary, and anchored in civic problems with authentic data.
- Technology (esp. CODAP + open data ecosystems) can equalize participation—age and gender differences were negligible in perceived learning.
- However, students may not automatically connect data practices to broader decision-making. To promote epistemic recognition, we need more explicit epistemic framing and reflection.
  - Include prompts/tasks that ask: *What decision does this analysis support? What claim is warranted by the data (and what are the limits)? How would different data/assumptions change the conclusion?*

Phase 2 were made refinements based on Phase 1 findings

- More time and mentoring between modules; shorter, localized materials and curated multilingual datasets.
- Platform usability and translations; embed surveys/reflective tools inside the platform for easier data collection.
- Make the "data → argument → decision/action" bridge visible (deliberative dialogue, explicit transfer prompts).

Co-funded by the European Union

18

Conclusions

Key takeaways and next steps

Bottom Line

- Engagement and data learning are strong; the next step is making **transfer to civic decision-making** visible to students.

Next iteration priorities

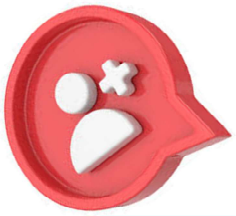
- Make the epistemic purpose explicit: *Why this analysis? What decision could it inform?*
- Build structured reflection: evidence → uncertainty → values → action
- Strengthen bridges to civic decision-making: prompts, rubrics, public communication tasks
- Support cycle-based refinement with teachers: CoP mentoring, curated datasets, local-language supports

Closing Line

- Data science literacy in STEAM becomes a **literacy practice** when students learn to **interpret, question, and act with data** in civic life.

Co-funded by the European Union

19



Follow Us

https://www.facebook.com/profile.php?id=61560406535213

https://datasc4ed.euc.ac.cy

https://www.instagram.com/datasced4cien/

Co-funded by the European Union

20