https://doi.org/10.53656/math2025-1-7-add

Educational Technologies
Образователни технологии

ADDRESSING DATA LITERACY IN INFORMATION SYSTEMS EDUCATION

Kalinka Kaloyanova[1,2)]

[1)] *Faculty of Mathematics and Informatics*
Sofia University St. Kliment Ohridski (Bulgaria)
[2)] *Institute of Mathematics and Informatics,*
Bulgarian Academy of Sciences (Bulgaria)

Abstract. Data literacy is essential in our data-driven society. The role of data in education is even more significant. Obtaining data literacy skills is important for both students and educators. Aligning data literacy in computing students' education will enhance students' learning experiences and better prepare them for future professional positions. This paper discusses the importance of data literacy in Information Systems undergraduate education, emphasizing the need for students in this field to understand and use data effectively and aims to give specific guidelines for including data literacy in information systems undergraduate education. A specific approach to incorporating data literacy components into one of the major courses in the IS2020 curriculum – the Systems Analysis and Design course is presented. The approach could be applied to other computing disciplines to enrich them with data literacy elements.

Keywords: data literacy; education; information systems curricula; system analysis and design

1. Introduction

Business decisions in our society are increasingly driven by data. To address the need for effective data processing and understanding, companies need employees who understand, interpret, and use data effectively. These are the major elements of data literacy (DL).

There are several definitions of data literacy but all of them notice data literacy is connected to the "ability to read, understand, analyze, and communicate with data". Gartner Glossary[1] defines data literacy as "the ability to read, write and communicate data in context, with an understanding of the data sources and constructs, analytical methods and techniques applied, and the ability to describe the use-case application and resulting business value or outcome".

Building a set of skills and competencies as the foundation of data literacy is becoming an important part of education, especially in universities.

This paper explores the necessity of integrating data literacy elements into computing curricula and more specifically in Information Systems (IS) undergraduate education. Reflecting some specifics of the IS programs in comparison with other computing programs, this article aims to give specific guidelines for including DL in IS undergraduate education. Given the challenges associated with this integration into the complex IS curricula (long time for discussion of different topics, large working groups of educators and professionals), the study reveals a possibility to involve data literacy elements in existing IS courses and encourage students to build essential data literacy skills during their education.

Based on one of the basic mandatory courses in the IS undergraduate programs – the Systems Analysis and Design course – it is shown how different aspects of data literacy could be outlined during the course.

2. Data literacy background

Data literacy is an important indicator of how well an individual or organization handles data to achieve a specific goal. It does not refer to simply viewing or scrolling information but to searching, understanding, analyzing, and making sense of data.

When considering individuals, for example, employees in an organization, a good level of DL would allow them to make better professional decisions based on data in their work. This means that they should not only have knowledge of how to find, read, and analyze data but also know how to use it to make the right professional decisions.

On the other hand, it is important for every organization to have employees who obtain a high level of data literacy. Moreover, in order to be more competitive, organizations should encourage their employees to increase their DL and create conditions for this to happen.

Depending on the circumstances, data literacy can encompass a broad or narrow set of skills and levels of knowledge.

In general, data literacy affects different aspects of working with data – understanding, analyzing, visualizing, and applying data to a task. Understanding data involves examining various data sets and assessing their quality. Working with data involves a lot of activities for creating, acquiring, cleaning, and managing data, and all these require technical knowledge. Analyzing data encompasses computational thinking, breaking down complex problems, applying logical reasoning, and using sophisticated tools. Visualizing data relates to grasping insights from data sets, recognizing patterns and relationships, and understanding how they apply meaningfully to the situation. Applying data to a task involves using data to convey narrative stories and support hypotheses and arguments (Knight 2024).

Data handling entails numerous tasks, including data creation, acquisition, cleaning, and management, all of which need technological expertise. Computational thinking, decomposing difficult issues, employing logical reasoning, and utilizing advanced tools are all necessary for data analysis to extract insights from big data volumes, identify trends and connections, and comprehend how these relate to the situation in a meaningful way.

Several studies (Inverarity et al. 2022) present more extensive research on the definitions of the DL and the competencies related to this concept have been studied. In (Pörtner et al. 2024) an assessment of competencies applicable across different industry sectors is presented. The authors present a list of more than twenty competencies, discussed in different sources. Most of them are, actually, skills, as the skills are seen as more practical abilities, that people acquired during their education and practice. Competencies, on the other side, are broader and present a combination of skills, knowledge, behavior, and attitude helping people to effectively perform a task or a role in a particular environment.

Further, the items on the list are separated into several broader categories (Pörtner et al. 2024): Data Foundations, Data Management Skills, Data Analysis Skills, Data Interpretation, Data Problem Solving, Communication Skills, Data Culture and Advanced Skills.

As can be seen, these groups aim to capture critical thinking abilities in addition to technical skills.

In the first category data definition, data sources, and data quality issues are covered. The next two groups talk about data processing technical and analytical concerns. All activities of data processing, from data collecting and cleansing to data analysis, data modeling, etc., are covered in the fourth group devoted to the need for accurate data interpretation. Computational abilities, which include dealing with programming languages, algorithms, and specialized data processing tools as well as data tools that facilitate data analysis, are also included in this category. The last important group concerns ethical considerations.

Building competence and skills for DL should begin at school. However, they are best developed at the university level. The universities are the right place for both teachers and students.

2.1. Data literacy in education

When exploring existing research on data literacy in the education field, it is noticed that many researchers pay attention to the development of DL skills by teachers themselves. In (Mandinach & Gunmmer 2016) the authors discuss the knowledge, as well as the skills and dispositions that teachers need to apply in their work.

Among the most discussed topics of data literacy in education are how to build the capacity of educators and students, how to implement technology and tools more effectively, and how to address issues related to ethics and privacy of data (Fagerlund et al. 2025).

Data literacy not only encompasses students' technical skills in working with data. Data literacy fosters understanding and interpreting data in different forms, critical thinking, communicating the story behind the data, and using data to be more efficient in their work. These competencies are also essential for the business, as today all companies strive to make informed business decisions.

Because teachers are expected to interpret data in different ways, starting with their students' performance data, it is important for them to improve their data literacy. The challenges here relate to both teachers' effective use of data themselves and creating or adapting programs that enhance their students' data literacy.

As for educational programs, they still leave a lot to be desired. In most cases, teaching focuses on small-scale data sets needed to solve specific tasks. While useful for solving a specific task, working with such data generally does not enable students to gain experience and skills for working with larger and more complex data sets used in real-world systems. This is very important for computing students.

2.2. Data literacy in the computing field

Most studies are focused generally on the common presentation of DL in school or university education (Lee et al. 2024). There are still no specific studies published on how to approach these issues more purposefully in the computing programs at the university. For students in the computing field, data literacy competencies and skills play an important role in their education and future careers. This group requires both foundational and advanced competencies in data processing, data analysis, and data communication, along with extensive technical and computational knowledge.

Speaking of competencies and skills, let us note again that while DL competencies represent broader categories or areas of expertise, encompassing the knowledge, abilities, and behaviors needed to work effectively with data, data literacy skills are more specific, actionable abilities within those competencies – practical things that a person can do to demonstrate proficiency.

The computing field encompasses several directions in undergraduate education such as computer science, software engineering, information systems, cybersecurity, etc. Data processing and analysis is a core activity in each of them. A good level of DL competencies and skills ensures that compu-

ting students can create data-driven solutions, applying critical thinking to effectively create software applications.

And yet there is a difference between different computing-related disciplines. While computing students focus more on technical implementation, IS students typically operate at the intersection of business and technology. Data literacy enables them to understand and leverage data within a business context, driving better decision-making and strategic planning.

Data literacy competencies and skills of information systems (IS) students are more tightly connected to their field, which connects technology, and business, and emphasizes the role of data-driven decision-making. IS students need a balanced set of technical, analytical, and business-oriented skills that help them manage data to solve organizational challenges.

2.3. Data literacy in Information Systems programs

Because of the specific of information systems as a special type of software product, the specialist in the information systems area should understand not only the technical aspects of data processing – data collection, data modeling, storing data into the database, data analysis, etc. They also need to understand the business context – the business goals of the organization and the role of the information and information system in accomplishing these goals. In many cases, the IS specialists provide the bridge between different teams in an organization – the technical and the business ones.

The latest IS curriculum – the IS2020 Competency Model[2] is a competency-based educational framework for undergraduate Information Systems (IS) programs. It is based on the competency model, highlighting both foundational knowledge and practical skills are important for IS professionals. Considering also that data and information management issues are crucial in the area, the intensive involvement of data literacy competencies and skills in IS education is highly relevant.

Introducing data literacy concepts, competencies, and skills in the existing courses, promoted by the IS2020 guidelines could be an effective way to help students gain the necessary competencies and skills. The courses on Database Management, Systems Analysis and Design, and Project Management, for example, provide a good base for data processing as many data aspects are discussed. Their content is appropriate for discussing major issues such as data understanding, data modeling, data analysis and visualization, effective data management, as well as data ethical and privacy considerations.

Certainly, the holistic approach of integrating data literacy across the curriculum should be followed, as every course could contribute to building DL competencies and skills. However, modifying the complete curriculum is a long and complex process. In the meantime, the incorporation of DL con-

cepts can start with one of the courses mentioned above, which more actively discusses data-related issues. In the following section, such a solution based on the Systems Analysis and Design course is presented.

3. Integrating data literacy in the System Analysis and Design course

Systems Analysis and Design (SAD) course is a mandatory course of the undergraduate IS programs. The course focuses on the early stages of software development. It starts with addressing business requirements and how they are collected and analyzed in order to clarify the functionality of the future system. The course further discusses how to design the system to meet the defined requirements.

During the course, students deal with a lot of data coming from different sources and apply various methods for collecting and processing data – first, for requirements gathering, and in the next stages – for exploring the information content of the system developed. They also apply data modeling using different models and techniques. Finally, an initial design of the system is developed.

From this perspective, the course is quite appropriate for integrating many DL competencies and skills.

The next sections demonstrate how this integration can be achieved.

3.1. System Analysis and Design course

The main goal of the SAD course is to teach students how a new IS should be developed by understanding user requirements and specifying the functionality and services the new system should provide. The course emphasizes the analysis and design phases of IS development.

Systems analysis and design deal with planning the development of information systems through understanding and specifying in detail what a system should do and how the components of the system should be implemented and work together. The role of system analysis is to resolve business problems by analyzing information system requirements and during the design phase – to create a blueprint of the new system that meets these requirements while applying a set of design methods and techniques. This course deals with the concepts, skills, methodologies, and techniques that extensively use data from different sources. In this way, DL components can be naturally integrated into the course in several directions.

Students need to learn how to align system functionality with organizational goals using all data identified during the process of requirements gathering. They also should identify specific data needs of the particular system.

Further, students use data modeling techniques and discuss data maintenance in the system design. This also considers defining data characteristics, including data privacy, data security, etc. Data quality and data consistency also should be accounted for.

How these elements can be integrated into a SAD course is discussed below based on the implementation of this course in the undergraduate Information Systems program at the Faculty of Mathematics and Informatics - Sofia University St Kliment Ohridski - *Information Systems Analysis and Design* course which is taught during the 7^{th} semester of student education.

In addition to the topics discussed above, the practical component of the course includes project development, which requires students to go through the steps of systems analysis and design to solve a real-life business problem. In this way, students consolidate the knowledge they have gained and acquire useful skills in the IS field.

3.2. SAD project and DL

The project objective is the develop a small information system. Students focus on the initial phases of the IS development and prepare a prototype of the system based on the requirements analyzed. The object-oriented methodology is followed for system development. Every team presents its vision for the goals of the information system, describes its functionality through a use case model, and presents an initial user interface (UI). Each of the tasks set for the projects practically affects DL elements.

Students work on projects grouped in teams. Every team consists of 8-10 students. At the end of the course, each team submits their team project and each student submits an individual project, which includes the results from the individual tasks he has worked on during the team project.

Student teams present the project results in 2 parts – one intermediate and one final. In both cases, students submit documentation and make a presentation. The goal of the intermediate submission is to provide an initial view of the system development. In the final version, students are required to present a set of artifacts (table 1) that describe the system project.

Students start work on the project with an initial vision of what the system should do. The task of the initial description of the functionality of the new IS understand the purpose, goals, and general structure of the system. It's based on foundational knowledge about how information systems work, including data types and system components.

DL elements included in this part concern:

- Aligning system objectives and functionality with business goals.
- Understanding different types of data (structured, unstructured, quantitative, qualitative) and their relevance to system functionality.

• Basic understanding of data flows within a system.

To define the functionality of the system, students need to identify all stakeholders. While preparing the list of stakeholders and functionality that meet their interests the students should understand each stakeholder's different roles and responsibilities and how their needs interact with the system's data – input data, reports, data analysis, etc. It must be ensured that all stakeholders understand data dependencies and flows. This information should be detailed and documented with a focus on the data:

• Stakeholder analysis with a focus on data needs and expectations.
• Mapping stakeholders to specific data requirements.
• Identifying how data needs shape system functionalities.

Once stakeholders and their interests are identified the methods for collecting their requirements should be chosen – e.g., surveys, interviews. The selection of methods for requirements gathering is an essential competency in any IS field. For this activity, students' curiosity can be fostered by encouraging students to ask "what" and "why" when analyzing business processes and exploring system inefficiencies.

After analyzing requirements, student teams present already formed and argued vision for the system they will develop. Here the focus on data is how the system will collect, manage, and analyze data in the future. To incorporate DL elements here, it is important to highlight how data handling aligns with user needs and include data scalability, accuracy, and accessibility as part of the product vision.

In the transition between analysis and design the use case modeling is fundamental for the new system development. Because the object-oriented approach is established for this course, to present system requirements students create a use case model, a complete list of use cases, and fully dressed descriptions of use cases.

Creating a complete list of use cases is a major task in designing systems. A detailed list of use cases helps map the system's functionality and its data interactions. Creating the use case model of the system presents how different users will use the system and gives a glance at how data will be used across the different components of the system. Describing scenarios of data flow steps on the use cases description presents the data interaction between a user and the system.

DL elements here concern several main tasks:

• Ensuring the list of all use cases reflects overall data management and interpretation including how data is presented to users.

• Describing data-related actions within use cases, such as data input, storage, processing, and output.

• Including detailed data requirements elements – input data validation, data transformation rules, data error handling, etc.

Together with functional requirements, non-functional requirements are discussed in the course, too. They include data-specific considerations like data scalability, performance, security and privacy issues, etc. Each team needs to define specific data-related non-functional requirements such as system response time under data load, how the system should scale with increasing data volumes, or compliance with regulations (for example, GDPR).

The task of building the system domain model is essential for integrating data in the newly developed system. The role of the domain model is to present the main conceptual data entities and their relationships. Also, business rules are captured in this model. In this way, the model presents real-world objects and concepts and can be easily discussed with users and stakeholders. The domain model is further detailed in UML class diagrams.

The technical part of the projects includes understanding the technology behind UML modeling diagrams (Sequence, Communication, Activity, State, and Class) and using tools (e.g., Microsoft Visio) to visualize system data, workflows, and interactions. UML modeling is included in the course because it can enhance students' understanding and application of system design as well as data principles. Between the UML diagrams used in the course, the class diagrams best match DL components. When creating class diagrams, students focus on defining data attributes for each class. Also how data is shared or linked in the system is revealed. As this part is widely discussed in other courses, e.g. Database Systems, it is only necessary to remind students of the principles of database design.

Creating the first elements of the user interface (UI), e.g. mockups, helps to apply system requirements in a tangible format and communicate them with the system users. It also integrates user needs and data flow into interface design. To effectively include DL components the mockups should:

• Display how data is presented and how users can interact with data using forms, reports, dashboards, etc.

• Demonstrate how data can be captured and searched in the system effectively.

• Reveal how data visualizations (charts, tables) will help users make decisions.

Specific guidelines for the inclusion of DL elements in some of the above-discussed tasks and corresponding artifacts are shown in table 1.

Table 1. DL integration in the SAD course projects

Project Task/Artifact	**Ensuring DL elements integration**	**DL competencies and skills**
Vision	Explain the importance of aligning system functionality with business goals and t Clarify the need for data identification during all activities	Data interpretation Critical thinking Data-driven decision making
List of Stakeholders and functionality that meet their interests	Ensure students map the data needs of each stakeholder and describe how data will be used by each (e.g., reports, alerts, visualizations).	Data collection Data analysis
Methods chosen for requirements gathering	Describe different methods for requirements gathering and their pros and cons.	Data collection Data analysis Data visualisation
Use case model	Emphasize the need for a complete presentation of all system users and their interactions with the system.	Data collection Data analysis Problem-solving
Fully dressed description of all use cases	Highlight the importance of having complete and consistent data for accurate preparation of the use case documentation. Pay attention to students to carefully describe all data flow steps, data validation, and user interaction.	Data design Data validation
Non-functional requirements	Point out the role of metrics for measuring non-functional requirements.	Data privacy Data security Data governance

Project Task/Artifact	**Ensuring DL elements integration**	**DL competencies and skills**
Domain model	Underline that the domain model must present major conceptual classes with basic attributes and relationships. Remind students to ensure the domain model accurately reflects data entities, data attributes, and data relationships.	Data modeling Data cleansing
UML class diagrams	Emphasize the role of defining data attributes and relationships for each class. Emphasize the data flow in the system.	Data modeling Data storage and access Data privacy Data tools
User Interface elements	Notice the importance of UI in presenting the data that users enter into the system. Encourage students to demonstrate how data visualizations could help decision-making.	Data visualization Data validation
Presentation	Insists on project presentation not only to describe how the system supports business goals but also how data is processed, what data models are, etc. Encourage preparing presentations that combine data visuals and narratives effectively.	Data visualization Data storytelling
Communication	Clarify the role of communications in teams as well as with users and stakeholders when discussing data and information.	Data sharing/ communication

During the course, students provide two presentations – on intermediate results and on the final ones. All team members take part in them. Presentations help the students to show the result of their work and to justify the decision made by the team.

On the other side, preparing a presentation is another way of collecting and organizing information about the project, showing how the particular team works with data sources, and what data and how this data will be used in the system they develop.

One such way of delivering the SAD course encourages collaborative learning by engaging students in group assignments that focus on data literacy components. In addition, the various discussions during the lectures and teamwork help students to enhance their communication skills. The tasks that students work on, presented in Table 1, are actually performed tasks implemented in real information systems development projects. The competencies and skills for collecting and analyzing information at different levels and with various methods, creating information models, communicating with users, and presenting information content used within the course are real experiences and help students build up skills necessary for working on real projects in the industry.

5. Conclusion

This paper highlights major issues in data literacy and how they could be addressed in an important field of higher education – Information systems bachelor programs. IS undergraduate students should prioritize data literacy skills because of their unique roles and focus. They work at the intersection of business and technology, requiring an understanding and use of data to make better business decisions.

The paper presents a way of including data literacy components in an existing course of the IS undergraduate programs – the Systems Analysis and Design course. This course is of particular importance in the IS undergraduate curriculum. As the course deals with different data sources that must be carefully considered and analyzed, it is very appropriate to address the challenge of data literacy integration IS students' education. This course discusses the early stages of system development, starting with business requirements and the first steps of designing a system to meet these requirements. DL elements can be included in these objectives by emphasizing the role of data in analyzing systems, designing solutions, and supporting decision-making.

As the topics discussed in the SAD course are common to many other disciplines in the computing field, the proposed approach could be adopted in another case, following the main principles, discussed in the paper.

Future research will be aimed at enriching the course with more DL elements and implementing the described approach to other academic disciplines.

Acknowledgments

This study is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project № BG-RRP-2.004-0008-C01.

NOTES

1. GARTNER GLOSSARY, 2024. Data Literacy. https://www.gartner.com/en/information-technology/glossary/data-literacy
2. IS2020, 2020. A Competency Model for Undergraduate Programs in Information Systems, https://www.acm.org/binaries/content/assets/education/curricula-recommendations/is2020.pdf

REFERENCES

FAGERLUND, J., PALSA, L., MERTALA, P., 2025. Exploration of domains of educational purpose in K-12 data literacy education research, *Educational Research Review*, vol. 46. doi: 10.1016/j.edurev.2024.100663

INVERARITY, C., TARRANT, D.D., FORREST, E., GREENWOOD, P., 2022. Towards Benchmarking Data Literacy", *Companion Proceedings of the Web Conference*, pp. 408 – 416. ACM.

KNIGHT, M., 2024. *What Is Data Literacy? Definition, Components, Uses.* https://www.dataversity.net/what-is-data-literacy/

LEE, J.,ALONZO, D., BESWICK, K., OO, CH.Z., ANSON, D., ABRIL, J., 2024. Data literacy of principals in K–12 school contexts: A systematic review. *Educational Research Review*, vol. 45.

MANDINACH, E.B., GUMMER, E.S., 2016. What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, vol. 60, pp. 366 – 376.

PÖRTNER L., RIEL, A., KLAASSEN, V., SEZGIN, D., KIEVITS, Y., 2024. Data Literacy Assessment – Measuring Data Literacy Competencies to Leverage Data-Driven Organizations, *Procedia CIRP*, vol. 128, pp. 78 – 83.

✉ **Dr. Kalinka Kaloyanova, Prof.**
ORCID iD: 0000-0003-0222-7607
Faculty of Mathematics and Informatics
Sofia University
5, James Bourchier Blvd.
1164 Sofia, Bulgaria
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Bl. 8, Acad. G. Bonchev St.
1113 Sofia, Bulgaria
E-mail: kkaloyanova@fmi.uni-sofia.bg