Journal on Interactive Systems, 2024, 15:1, doi: 10.5753/jis.2024.3251

Human-Data Interaction Syllabus for Undergraduate and Graduate Courses

Thiago A. Coleti [**Universidade Estadual do Norte do Parana** | ***thiago.coleti@uenp.edu.br***]
Marcelo Morandini [**Universidade de São Paulo** | ***m.morandini@usp.br***]
Lucia Vilela Leite Filgueiras [**Universidade de São Paulo**| ***lfilguei@usp.br***]

State University of Northern Paraná, Rod. BR-369, Km54, Vila Maria, Bandeirantes, BR, 86360-000, Brazil.

Received: 15 March 2023 • **Accepted:** 25 October 2023 • **Published:** 01 January 2024

Abstract The phenomenon of the data deluge is a reality and the volume of data produced by people and companies is much greater than what can be handled and analyzed. Data play a crucial role in guiding the efficient utilization of technological resources for companies, aiding them in product and service management. Moreover, individuals who have become adept data producers and consumers are increasingly orienting their lives toward data. To address this evolving trend, there is a growing imperative to educate Computing professionals. These professionals are required to design technology solutions that facilitate the synergy between individuals and data, a phenomenon known as Human-Data Interaction (HDI). This paper introduces a suggested minimum syllabus for HDI courses, with the aim of addressing the key themes associated with the interaction between individuals and data. The complexity and depth of HDI topics justify a dedicated course, preventing the risk of essential content being fragmented or inadequately covered if dispersed across different courses.

Keywords: Human-Data Interaction, Data Science, Data Literacy, Information Visualization

1 Introduction

The phenomenon of data deluge, feared in the past decade, is now a reality: there is more data than institutions are able to handle. Accenture [2021] provided a recent consultancy report that showed that the ability to effectively use data in the corporate environment greatly depends on the existence of a chain of professionals able to understand, question the significance and take specific actions based on data analysis.

This chain is made up of data users as consumers, of data scientists as producers and intermediaries staff responsible for analyzing and discovering the insights. According to Hornung *et al*. [2015], this social construction highlights the need for Human-Data Interaction (HDI) designers within the organizations.

The concept of Human-Data Interaction (HDI), as outlined by Mortier *et al*. [2016], centers around placing humans at the heart of data flows. This involves creating mechanisms that enable individuals to engage directly with these systems and the underlying data. As a result, the role of HDI designers extends beyond the realm of computational, statistical, and machine learning methods. It encompasses a broader scope, including the integration of human-centric elements, addressing user requirements pertaining to data and information, ensuring privacy and transparency, employing effective information visualization techniques, fostering data literacy, and adopting methodologies that prioritize user-centered data design.

Hartson and Pyla [2012] describe human-centric elements as the aspects, features, or considerations within a design or system that are focused on meeting the needs, preferences, behaviors, and experiences of human users. These elements are aimed at creating solutions that enhance usability, user satisfaction, and overall effectiveness by taking into account the human aspect of interaction. In the context of HDI, human-centric elements might include factors like intuitive user interfaces, personalized experiences, accessibility features, ethical considerations, and a strong emphasis on user feedback and engagement.

Preparing the future computing professionals to work in this scenario is necessary and urgent. For Arass and Souissi [2018], data professionals are expected to have skills that can comprise the entire data life cycle, from its conception to its analysis, use and disposal. They are also expected to combine data skills with those from Human-Computer Interaction (HCI) discipline in order to place humans as the focus of design. The ACM Data Science Task Force's Computing Competencies for Undergraduate Data Science Curricula ACM Data Science Task Force [2021] declares HCI competences in the Knowledge Area Analysis and Presentation.

HCI competences are typically integrated into Computing curricula, including fields like Computer Science, Information Systems, Computer Engineering, and Software Engineering. In essence, this knowledge can serve as an enabling factor for Computing graduates seeking opportunities in the HDI field. However, the ACM Data Science Knowledge Area suggests that there is substantial content, both in terms of breadth and depth, to support the necessity of a dedicated HDI course. Still, this need is not yet a consequence of employment figures, but in the anticipated growth of the profession. A snapshot of the professional market, obtained from the Glassdoor[1] website (as of August 22, 2023), illustrates the number of job vacancies in various countries worldwide. It is noteworthy that there is demand for specialists in one of the most relevant aspects of Human-Data Interaction

[1] https://www.glassdoor.com

(HDI), which is information visualization, a number that can be compared to the vacancies for a more generic position in the field of data science. It is worth noting that in some countries, the ratio is 1:10, while in others, it reaches 4:10.

So, it can be really significant to prepare students to work in an area that includes human aspects, data-people relationship, artificial intelligence, information visualization, HDI tools and methods. In addition to the technical aspects, Data Literacy should also be considered, which Taibi *et al*. [2021] describe as the knowledge acquired by people that allow them to work with the available data and information in a useful, relevant and efficient way.

Thus, this paper presents a proposal of a minimum curriculum for teaching HDI in undergraduate and graduate courses (Lato or Stricto Sensu), lectures and workshops. For undergraduate and graduate courses, a discipline approach is proposed that contemplates, in more details, several aspects of the HDI. For the discussion of HDI in workshops and lectures, the goal is to deliver the explanation contents in a simplified way, but that can provide the minimum knowledge necessary to encourage students engagement with aspects of HDI in their future professional and teaching activities. The need to prepare students to work in HDI field aims to address both market and academic demands, such as:

1. The typical profile of a data professional is often associated with a Data Scientist and their skills, as mentioned in Ismail and Zainal Abidin [2016], Lopes *et al*. [2017], and Peek [2023], tend to focus on building statistical models, machine learning models, statistical analysis, and database manipulation. These aspects are predominantly technical and computational. Consequently, human factors might not be adequately considered or might be treated marginally, affecting the way data strategies are employed for and by people;
2. The increasing demand for skilled professionals to work with data is evident in recent research, such as the one presented by the Fordes Journal in Janssen [2022] and the School of Engineering of Tufts University[2], which highlight significant job market growth. According to the U.S. Bureau of Labor Statistics, the growth rate is projected to be nearly 28%;
3. There has been a substantial rise in the number of individuals using data-manipulating software applications that guide their actions and decision-making based on the information provided and secured by these applications (United Nations Development Group, 2017) . Mortier *et al*. [2016] emphasize that people now live in a data-driven world, and Maus [2015] further underscores the near impossibility of interacting with a software application without having personal data collected.

Undergraduate and postgraduate courses in Brazil lack a consolidated strategy for teaching HDI, with this content often being part of a larger discipline or not addressed at all.

This paper is an extension of the of the ”Insertion of Human-Data Interaction Contents in the Human-Computer Interaction (HCI) Discipline”, produced by Coleti *et al*. [2022] and published in the Annals of the XXX Workshop on Computing Education.

The organization of this paper comprises, in addition to this Introduction Section, a Section 2 - Theoretical Foundation, in which subjects that support this paper are discussed; the Section 3, which discusses how HDI and privacy issues were inserted in Computing Courses subjects; a Section 3 presents the actions taken to select and apply HDI and privacy content within the HCI discipline in Computing courses; Section 4 presents a proposal for the Human-Data Interaction (HDI) discipline for undergraduate and for stricto or lato sensu postgraduate courses in the area of Computer Science, as well as suggestions for lectures and dissemination workshops; Section 5 presents a proposition for discussing HDI through lectures and workshops; and Section 6, the conclusion of this paper.

Next section presents the Theoretical Foundation of this paper.

2 Background

This section presents the theoretical foundation of the contents that supported this paper, which includes the following subjects: a general overview of HDI, HDI Methods/Techniques and Data Literacy.

2.1 Human-Data Interaction

Haddadi *et al*. [2015] explain that Human-Data Interaction (HDI) studies phenomena related to the interaction of people with software and hardware applications that handle data and the consequent impact on people's lives. Hornung *et al*. [2015] highlight that studies in this area advanced as the volume of data produced and used became gigantic and started to influence the lives of people, companies, schools and other organizations, making society strongly data-oriented.

This field is also described by other authors as:

- Elmqvist [2011] defines HDI as the human handle, analysis and decision making in large, complex and unstructured databases;
- Cafaro [2012] defines HDI as the delivery of personalized data, within context and understandable for users;
- For Mashhadi *et al*. [2014], HDI provides support for features to provide individuals with access and understanding of data and how data affects their behavior;
- McAuley *et al*. [2011] describe HDI as the ability to manage several different personal databases and enable the user to control events using the data.

Although with certain differences, these definitions have some common elements that allow us to assume that HDI descends from the areas of Data Science (Big Data, in general) and HCI. This is due to the fact that, in Data Science, as presented by Amaral [2016], the entire process of data manipulation occurs in order to achieve a certain objective, through methods, techniques, tools and resources that allow data to be collected, processed, transformed, analyzed, disclosed and shared.

[2]https://onlinesoe.tufts.edu/blog/fixing-the-data-science-shortage/

The HCI area inheritance occurs with the phenomena of computer-intermediated interaction between people and data/information, since they participate in the manipulation of data as producers (source) and/or as information consumers, as discussed by Hornung *et al*. [2015]. Therefore, Evequoz and Lalanne [2006] present that it has become more common for people to use maps, graphs, images, interactive information visualization resources and intelligent automation systems to guide their daily tasks.

In this approach, hundreds of projects meet the characteristics of HDI can be identified, such as those presented below:

- **GPS applications**: researches similar to the one presented by Beber *et al*. [2017] analyze the use of location data to generate information about observed agents, which can be people, objects or animals (biodiversity monitoring) that are used by companies and people;
- **Social Networks**: Lebo *et al*. [2016] explain that through Data Analytics tools and Machine Learning algorithms, it understand the behavior profile of their users to direct content, products and services that may be of interest to their consumption.
- **Health monitoring tools**: can be defined as an approach known as Personal Informatics. As presented by Choe *et al*. [2014], it aims to present information about the individual who produces the data. Smart watches and other sensors collect physiological data from the person and specific algorithms transform them into information about the individual's health condition;
- **Data immersion applications**: exemplifying the approach proposed by Elmqvist [2011], data interaction approaches should allow the user to be inserted into the data environment and become part of it. This must be considered as not only as someone who visualizes the information, but who interacts with and interferes in them. For example, museums that the user can interact with, explore and delve into information through the use of tangible 3D objects;
- **Information visualization and exploration**: applications using maps and other visual resources that allow users to interactively explore information, as exemplified in the work of Cham *et al*. [2022].
- **Smart Home**: Basarudin *et al*. [2017] explain that this approach promotes personalized and user-focused experiences through the processing of personal data, which is manipulated by algorithms, *Machine Learning* techniques, Data Mining, Internet and Bluetooth, among other computing or electronics resources.

HDI field is relatively new. HDI research dates from the beginning of the 2010s, although the concern with issues of privacy and data security are older, exemplifying Lieshout and Kool [2008]. In the bibliography, we found specific papers that discussed the challenges of HDI. Specifically, we highlight the ones of Hornung *et al*. [2015], Mashhadi *et al*. [2014], Mortier *et al*. [2016] and United Nations Development Group [2017] that points about challenges in personal data usage.

These papers describe challenges such as: improvements and adaptations in Human-Computer Interaction (HCI) projects, user interaction and immersion with the data instead of just accessing the data; ownership of personal data, effective strategies for viewing information and resources for individuals to understand and control the handling of personal data and ensure their privacy.

Regarding the topics mentioned, some already receive the researchers, developers and governments special attention. This is the case of data ownership, information visualization and improvements and adaptations in HCI projects. However, these issues are strongly linked to data-based systems and their manipulation by people, which allows us to assume that the evolution of these systems and changes in people's data and information needs will lead to improvements needs.

Other subjects still require improvements in the context of technological and scientific developments, such as: the immersion of people in data manipulation and, mainly, in aspects related to privacy, security, freedom and other aspects of handling personal data. Aspects related to Software Engineering, Databases and HCI have also received scientific attention, as mentioned by Gorton *et al*. [2016], Ji *et al*. [2020] and Pekala [2017], since the construction of these applications may require adaptations of the methods and techniques used.

In the area of HCI, researches aim to improve the interaction of individuals with information, in particular issues related to understanding how their data are manipulated and to ensure their privacy is respected. Baig [2020] describes Privacy UX as the ability to develop the applications user interface to provide an adequate user experience for data subjects manage and ensure that your privacy is not affected. For Baig [2020], Privacy UX designers must concern with requirements such as: (1) Let Users Have Both Options: "Accept" and "Decline"; (2) Don't Be Too Quick to Ask for Consent; and (3) Keep the Human Factor Alive, people's interaction with applications can be more efficient and reliable.

Considering HDI and other products or services that handle data, Cavoukian [2020] presents the *Privacy by Design* (PbD), that aims to ensure privacy as an inherent, non-excluding requirement of all product or service development stages. Cavoukian [2020] highlights that PbD can be applied to different types of product and service projects such as: Information Systems, in particular HCI, Audit, Business or Infrastructure projects.

Finally, the endless possibilities for handling personal data lead to the need for *fair use* of data by controllers, concern for privacy, security and freedom of individuals and the possibility of providing means for people to interact with data to achieve specific goals, whether personal or professional.

Next section presents a discussion of some Methods and Approaches aimed at HDI.

2.2 Information Visualization

In this section, the Information Visualization (Infovis) concepts and approaches related to the need to build visual models for users and data subjects will be briefly discussed.

Sindiy *et al*. [2013] describes Infovis as an intersection between the areas of computer graphics, graphic design and interaction design, which seeks to build visual representations for data. Few [2016] adds that Infovis seeks to communicate and organize abstract data such as qualitative analysis, statis-

tical data, distribution of occurrences. To achieve these goals, can be used visual means such as texts, graphs, tables, maps and images, in order to allow the communicated message to make sense to your receiver.

Few [2016] explains that a *successful visualization is related to the degree that it encodes information in a manner that our eyes can discern and our brains can understand*. To achieve this, the following aspects in the design of an Infovis should be considered: (1) indication of the relationship between values and design variables; (2) accurate representation of quantities; (3) ease for handling of comparison, ranking and organization of information; (4) make it obvious how information is to be used.

Hsieh [2016] and Few [2016] also highlight that information visualization is supported by aspects of cognitive psychology related to human perception. From this concept, the *Gestalt principles* are inherited, which guide the accuracy of descriptions of human visual behavior. Some Gestald Principles presented by Few [2016] are: Proximity, Similarity, Enclouse, Closure and Connection.

Knaflic [2019] explains that Infovis has become a fundamental element of the data manipulation process, since, whether applied effectively, it can mean the difference between success and failure for communicating findings to a specific audience. Thus, Infovis should not be limited to presenting data in graphs or drawings, but the designer should think about how to tell and/or communicate the story to his target audience.

Infovis has always been the focus of Data Science projects for data experts. However, Huang *et al*. [2015] points out that the avalanche of applications aimed at people (not specialists) has allowed a greater interaction of people with information and due to this, strategies for Personal Infovis have been developed by researchers and developers.

Huang *et al*. [2015] explain **Personal Infovis** as the design of visual representations and analytical approaches for the use of people in a personal context. The Personal Infovis challenges are the different perspectives, objectives and understanding of the same information in a completely different way. Thus, providing new perspectives for users to analyze and explore information is a great challenge for designers.

The different goals and information needs can be exemplified by the following Personal Infovis approaches:

- **Personal Visualization**: described by Huang *et al*. [2015] as the ability to provide interactive features for individuals to explore data in a personal context of use without necessarily performing complex or computer-supported analyses;
- **Personal Visualization Analytics (PVA)**: Huang *et al*. [2015] explain that, in this approach, complex analyzes supported by computers are added to the visualization of information. The purpose of PVA is not only to explore a dataset, but to analyze it more in-depth way in order to search for better solutions for the individual. This strategy uses interactive resources for customizing query parameters and adjusting them to the desired information needs. Cafaro [2012] explains that this approach requires the user to have prior knowledge of the elements that influence the final information, in order to manipulate it effectively.
- **Embodied Interaction**: described by Elmqvist [2011] as a more complex strategy, used to interact with highly dynamic information in which the individual's behavior or needs interfere with the final result. An example of this type of interaction can be an interactive Museum, in which the person has the possibility to interact with the content.

In addition to previously presented approaches, strategies aimed at presenting information about how personal data are handled, for the purposes of Transparency and for Privacy can be presented in the course. These approaches make use of traditional design components to provide interfaces and information able to communicate about events performed on personal data, for example:

- Holtz *et al*. [2011] and Efroni *et al*. [2019] discussed the development and use of icons with designs created specifically for Privacy and Transparency. Their discussions involve the advantages and concerns of communicating privacy aspects with the use of icons, since it can facilitate understanding events and agents involved in the use of personal data or confuse people even more;
- Cradock *et al*. [2017] and Murmann and Fischer-Hübner [2017] presented strategies for Transparency of Personal Data such as the use of Maps, Colors and Markers to help display the location of operators, controllers, data servers and other locations of interest to the individual and also to emphasize the importance and the sensitivity of information besides the events performed with the data;

Finally, Infovis is a fundamental HDI component, due to the fact that people's interaction with data must take place through audio-visual elements in order to facilitate and improve communication and understanding of information. Knaflic [2019] writes that the success of a good Infovis is related to understanding the context of use of data and information, understanding the target audience and choosing an effective model for presenting information.

Next section presents Data Literacy concepts.

2.3 Data Literacy

A third subject related to HDI that has aroused the interest of the academic and industry community is Data Literacy. According to Amo-filva *et al*. [2022], Data Literacy is the knowledge necessary to read, analyze and transform raw data into relevant analytical information, in addition to analyzing and work with possible achieved results. Actions aimed at Data Literacy can be seen in companies, as described in the report by Barth [2022], and also in academia, by Vivacqua *et al*. [2019] and Taibi *et al*. [2021].

Barth [2022] highlights that the global pandemic was a determining factor for organizations and people to realize the need to rely on data to make their decisions. Therefore, the concern with teaching how to work with data gained attention from companies and educational institutions. Projects such as those presented by Taibi *et al*. [2021] and Amo-filva *et al*. [2022] highlight initiatives to apply data teaching and

also to produce analytical tools to support teachers and students. In the DataONE Project, Allard [2012] it is pointed out that Data Literacy is essential for future scientists to work efficiently and productively with data, and that Data Literacy should be the basis of teaching in all stages of the data life cycle.

Barth [2022] also presents seven Data Literacy principles, which according to him, all employees of a company must be aware of and apply to support the organization to be truly data-oriented. These principles are:

1. Foster a culture of humility and curiosity: Data Literacy is not just about data, but refers to the development of soft-skills such as curiosity, creativity and collaboration;
2. Encourage employees to put training into practice: encourage people to seek training to work with data;
3. Bring everyone on the journey: all people involved in the organization, research etc should be engaged in using the data to get the best decisions;
4. Focus on the desired outcomes: clearly identify which problem can be solved with the data;
5. Measure the impact of the efforts: analyze whether the efforts applied to Data Literacy have an adequate return on investment for the organization in a specific period of time;
6. Adopt a systemic perspective: seek that Data Literacy is an institutional process implemented as part of the organizational culture, so that all sectors and departments act together to guarantee the culture of data usaget;
7. Decide what technology can meet the business needs: select, evaluate and monitor the shitty technologies, to ensure that the best and/or most appropriate ones are applied in your organization.

This paper also shows the concern in teaching people to work with data, within the context of organizations. In academia, teaching projects focused on Data Literacy are already applied, exemplifying the projects below Barth [2022]

Taibi *et al*. [2021] presents the DEDALUS Project as an EuropeanFunded project in the framework of the Erasmus+ Strategic Partnership in Higher Education action. The aim of the DEDALUS project is to equip students with the necessary competences to cope with future digital challenges and to create an additional value for the enterprises where they will be employed. DEDALUS arose from the need of European companies for qualified professionals to work and exploit the maximum data. This project is based on the following skills: (i) Handling digital information formats in professional Higher Education contexts in general; (ii) Data collection and Interpretation in Research, Development and Teaching in Higher Education; (iii) Critical Thinking; (iv)Specialist Data processing in Information Technology study fields; (v) Data Management. DEDALUS was performed at Universities of Vilnius, NoviSad and Southampton as pilot projects.

In Amo-filva *et al*. [2022], a tool that analyzes Logs of use for the Moodle environemt and translates this information into Infovis approaches was evaluated. This evaluation considered how much the information presented was appropriate for the professors' and how much it facilitated to immerse them in decision-making based on data. As a result, the authors believe that this tool still has some limitations, but it is efficient in producing information and helping teachers in the analysis of students' interaction.

In Vivacqua *et al*. [2019], the authors discuss the implementation of Data Literacy in primary and secondary schools. The authors' justification lies in the fact that schools need to train people with critical, analytical and creative capacity to study and make decisions based on data. An application strategy in schools is presented in this paper and the authors also highlighted the possibility and/or need for support from teachers, who are also expected to be interested in data training.

Regarding the courses related to the HDI, an exploratory research was carried out aiming to verify their existence as well as their academic profile regarding the construction of knowledge and skills on the HDI. There are ongoing HDI-related courses and research programs, such as: (1) the Data Interaction and Visual Analytics course at the Rutgers School of Science and Arts [3]; (2) the Systems for Human-Data Interaction [4] course at Comumbia University Coms; and (3) the Human-Data Interaction [5] research program at Swinburne University of Technology.

Considering the approaches and programs presented above, it can be said that there is a movement towards promoting the training of professionals and non-professionals in working with data systems. Strategies include both preparing technology developers, data analysts and also people to use data and make effective data-driven decisions.

Although with particularities in each proposal, the idea that knowing how to work with data is fundamental for the development and use of future products and services.

Next section presents the experience of applying HDI and data privacy concepts within HCI courses.

3 Human-Data Interaction into Human-Computer Interaction Courses

This section presents the actions taken to select and apply HDI and privacy content within the Human-Computer Interaction (HCI) discipline in Computing courses. This section refers to the paper **Insertion of Human-Data Interaction contents in the Human-Computer Interaction (HCI) course**, by Coleti *et al*. [2022], published in the Annals of the XXX Workshop on Computing Education.

The subject of Human-Computer Interaction, in computing courses in Brazil, was proposed by Boscarioli *et al*. [2012], and has a base curriculum focused on human, social and technical aspects of people's interaction with computational tools.

In HCI´s context, Mortier *et al*. [2016] explain that the increased use of software applications has led to concerns

[3] https://www.cs.rutgers.edu/academics /graduate/m-s-program/course-synopses/course-details/16-198-526-data-interaction-and-visual-analytics

[4] https://columbiaviz.github.io/2020s_w6998/

[5] https://www.swinburne.edu.au/research/institutes/data-science/human-data-interaction/

about the users´ privacy, since it is usual that users provide personal data allowing the identification of the person, their actions, customs or preferences. Thus, according to Toledo [2020], the need to deal with the privacy of software users grows as the manipulation of personal data by technological resources increases and interferes more actively in people's lives.

Therefore, it is assumed that it is significantly relevant to provide learning mechanisms for future HCI designers and developers to consider questions about the handling of personal data related to their projects. Among these issues, Haddadi *et al*. [2015] highlight: transparency of actions; accessibility and manageability of data usage permission; readability and privacy . So, it is necessary to discuss aspects of Human-Data Interaction (HDI) and privacy within the syllabus of HCI disciplines.

The organization of contents within the HCI discipline had been following the principles of: (1) not to change the already consolidated syllabus of the HCI disciplines, but insert concepts of HDI and privacy to complement and enrich the discussion of the content; and (2) because it is a relatively new area, the reference bibliography, in the form of textbooks, is still scarce, so the proposition of technical/scientific articles and regulations for the use of personal data as bibliography would be necessary; and (3) the execution of the contents was strongly supported by privacy issues and the impact of handling personal data on users.

3.1 Human-Computer Interaction Course

Human-Computer Interaction (HCI) field of study, as Rocha and Baranauskas [2003] explains, studies the interaction phenomena among people and computers. Benyon [2011] and Cybis *et al*. [2015] highlight that with the popularization of highly interactive software applications, the need to consider human aspects in systems design, to allow learning and using available resources became a necessity for software projects, in particular: the HCI projects.

Cybis *et al*. [2015] and Rogers *et al*. [2013] highlight actions such as: identifying users' needs and characteristics, developing design and user experience and evaluating the application's usability contemplate the range of tasks carried out by HCI professionals. As a subject in computing courses, Lopes *et al*. [2016] emphasize that HCI comprises a mandatory curricular component in the vast majority of courses in Brazil, or at least, HCI is present as optional or elective.

The syllabus of the HCI in computing courses was proposed in 2007 by Boscarioli *et al*. [2012]. It is being rediscussed and updated in editions of the Brazilian Workshop on Education in HCI (WEI) in 2013 and 2014 . In their work, Boscarioli *et al*. [2012] show that, although the contents may present specific differences to meet the profiles of Computer Science (CC), Information Systems (IS) and Computer Engineering (CE) courses, the main topics are:

- **Introduction to HCI:** evolution, areas and discipline, interface and interaction, usability and accessibility;
- **Theoretical foundations:** human factors, Ergonomics, Cognitive Engineering and Semiotic Engineering;
- **HCI Assessment:** User observation and monitoring, Interpretive assessment and Predictive assessment;
- **Human-Computer Interaction Design:** Interaction Styles and Guidelines, Design Guidelines and Patterns;
- **Process of Design in HCI:** Elicitation, Interaction Modeling, *Storyboarding* and Prototyping;
- **Domains and Platforms:** HCI in Games, Collaborative Systems and Mobile Devices.

The proposed contents promote the construction of the student's knowledge, since they provide fundamental theoretical knowledge and lead to technical training, with focus on the intercface´s design and evaluation. They also consider social and organizational aspects. In the content of the HCI discipline, some themes require attention and constant updating by the instructors (lecturers), as they are directly related to the hardware and software platforms, which might require constant updates, such as interaction approaches (desktop, Web , Web-Responsive, Mobile), as explained by Wroblewski [2011] and Neil [2014].

Other topics may require less periodic updates, but require the instructor's attention. For example, the concept of accessibility might changed during the time and the resources available are constantly improving. Another example is the set of Usability attributes, initially defined by ISO-9241 and later updated by ISO-25010 that still are in continuous evolution.

Finally, it should be noted that the HCI discipline prepares the students for one of the greatest market demands within the computing area and the impact of its products for the final user places the discipline as one of the most relevant in the computing course.

Next section presents the process for including the topic of privacy in HCI disciplines in computing courses.

3.2 Content Selection

This subsection describes how HDI content was selected for discussion within HCI disciplines. The contents were organized to complement existing contents in the HCI discipline, as described below.

3.2.1 Introduction to HCI

For the initial stage of the HCI discipline, it was decided to include the discussion of the fact that the people (users) interaction with computational resources provides an environment for data manipulation by people and companies. Thus, discussions are proposed on:

- Maus [2015] consider that it is practically impossible to interact with a computer without producing or consuming data;
- related to the HCI projects, constant concertns´ impacts, or even the requirement, of handling personal data and the need for data privacy;
- the ability of the person to produce and interact with personal data and information produced by him or herself or by other people;

- the ecosystem created by the massive production, manipulation and consumption of user data and information and how it affects people's lives.

As a HDI´s introductory bibliography, it is recommended to consider as main papers the following ones: Bellamy and Alonso [2016] Mortier *et al*. [2016] and Murmann and Fischer-Hübner [2017]. These papers discuss the general approach to handling personal data and present real examples of the environment created by handling personal data. It is assumed that this bibliography allows the student to create an initial mental model about the relationship between people's interaction with technology resources and the impact of handling personal data and data privacy.

3.2.2 HCI Design and HCI Design Process

Regarding the conduction of HCI projects and the use of style guides design standards, it was assumed that the contents on HDI and privacy should be of a practical nature in order to prepare the students to deal with elements of interface and with the interaction that can support data manipulation and privacy in the software application. Therefore, the suggested contents are:

- *Privacy by Design*: define privacy as an inherent, inseparable and priority part of projects, so that it must be addressed from the initial stages to project delivery;
- *Quality of Information*: proposed considering the need for personal data transparency, which is characterized by the ability to inform the user about the personal data handling and the actions related to their privacy in an objective and relevant manner. At this point, issues such as readability, quality of information and text quality must be addressed to prepare the students to abstract the techniques that handle personal data into accessible and understandable information for people.
- *TR-Model*: still in relation to transparency, the TR-Model is a model that describes a minimum set of information to be presented to the data subject and how the information must be presented from the point of view of design and readability. It was assumed that the use of TR-Model can help students to understand the context of transparency regarding what information to present and how to present it, thus promoting a relationship with the previous theme of Information Quality.

For the bibliography of this stage, it is suggested to study the following paper: Coleti *et al*. [2020] as it presents the TR-Model itself. Also, Hosseini *et al*. [2016] and Kumar and Jakhar [2010] should be considered as they discuss concepts of transparency and privacy in projects. Lee *et al*. [2002] and Kandari *et al*. [2011] should also be studied, as they deal with information quality concepts and techniques.

3.2.3 HCI Assessment

For HCI assessment tests, it is assumed that there is the need for discussion with the students to consider local and international regulatory texts such as the brazilian *Lei Geral de Proteção de Dados* (LGPD) and the european General Data Protection Regulation (GPDR) in addition to that practices that planned to deal with the applications that consider privacy.

Models, such as or presented by Tom *et al*. [2018] and Santos *et al*. [2018], in addition to TR-Model, can be used together with already consolidated evaluation techniques, inspiring heuristics, supporting inspections and usability tests and providing guidelines, tips practices, *checklists* or the same metrics for the evaluations.

Therefore, the quality of information and the (*Privacy by Design)* must be strengthened as resources for this stage of the project.

3.2.4 Domains and Platforms

Since HCI and HDI can encompass several different areas, it is suggested that the aforementioned topics be discussed using specific areas such as games, information Visualization, mobile devices, natural user interfaces, smart toys among others. These areas, undoubtedly, can generate massive actions for the manipulation of personal data and consequent impact on the users´ privacy.

Finally, it is strongly recommended to discuss the manipulation of sensitive data for tools such as software applications for health or safety area. These may handle with data that can critically interfere in the lives of their users.

Next section presents two case studies carried out with the proposal of HDI content and privacy in HCI disciplines.

3.3 Case Studies

This section presents two case studies of the application of HDI and privacy contents in HCI disciplines for computing courses. To describe the case studies, we considered the recommendations presented by McGill *et al*. [2018].

3.3.1 Content offered in a lecture

The first experience was conducted in the form of lectures in HCI disciplines. The lectures were held in person, in 2018 and 2019, with approximately a hundred and twenty students that included: a group from the Computer Engineering course, at the Polytechnic School of the University of São Paulo (POLI-USP), during the daytime; and two classes from the Information Systems course at the School of Arts, Sciences and Humanities of the University of São Paulo (EACH-USP), in the evening period.

The lectures were conducted as part of an HCI course within a Master of Business Administration in Data Science program. The course was led by a seasoned professor with a remarkable 35 years of experience in the field of Human-Computer Interaction (HCI). The lecture itself was delivered by a doctoral student, aged 39, who boasted a decade of professional experience in both industry and academia specializing in HCI. This student was concurrently pursuing research in the realm of Human-Data Interaction (HDI), focusing specifically on the theme of Personal Data Transparency.

Furthermore, the lecture received substantial support from a second professor who was not only a doctoral student supervisor but also a prominent figure in the HCI domain. This professor held an extensive background of over 20 years in

academia, further enriching the lecture's depth and quality. Notably, it is worth highlighting that the instructors mentioned are also the authors of the paper in question.

We did not collect demographic data specifically, which does not allow us to accurately quantify each student's profile. However, some general information was obtained, which allows us to state:

- All students were graduates, once the course was an MBA/*lato sensu* postgraduate course. Thus, we assumed their age as, at least 21 years old;
- It was a mixed class, made up of male and female students;
- The students worked in different areas such as Computing, Marketing, Law, Administration among others and, although a course with a Computing profile, their interests in working with data motivated their participation in the course;
- It was assumed that the participants had certain financial condition and did not belong to a group of financial vulnerability, as the cost of participating in the course was relatively high. However, details of this information were not requested at the time.

It was known that the discussion about the contents in a lecture would not be effective for the design project aspects, as it can be in a complete course, however it was assumed that it would be possible to highlight the importance of privacy and other aspects of HDI in projects of design and arouse the students' concern for the subject.

The lectures aimed to contextualize HDI into HCI aspects in order to arise the students' concern with this are. It were presented under the title: Human-Data Interaction: Reality, challenges, concerns and expectations regarding the use of personal data. They lasted approximately sixty minutes each for three different classes and their contents were: (1) the ecosystem for the use of personal data; (2) transparency of personal data; (3) data protection regulation; and (4) privacy by design. We used a Power Point presentation to conduct an interactive lectures that lasted 60 minutes and the students were allowed to ask questions during the lecture, but the students did not do it and left the questions for the end.

At the end of each lecture, the students were invited to participate in an activity in which they had to choose mobile applications or websites to analyze, from the HDI point of view, the aspects of handling their personal data. The following tasks were suggested: registering, reading the privacy and security policy, consulting access histories and analyzing possible directions for products and services.

A discussion was held on relevant and/or worrying points that were identified by the students, which stood out: [6]

- *"..the application already had my registration data, but I never used the physical store or any virtual service, how is the data here??"*;
- *"... in the privacy policy, I am authorizing you to do anything with my data..."*;
- *"...my life is in the hands of the application managers and I can't do much..."*
- *"... I wouldn't have subscribed to this app if I'd read the policy first, but it's impossible to read so much contents written for judges and lawyers.."*;
- *"...the app won't let, or at least won't make it easier for me to change my permission or file a complaint..."*.

These reports allowed an initial view that led users to the following emotions: (1) surprise, as there were cases in which the participants realized that the applications already had their personal data or manipulated them in a way that was not well perceived or known; (2) concern, as the handling of personal data seemed to go beyond what was considered a limit established by each of them; (3) discouragement, since the difficulty or inability to act in case of inappropriate use of data was clear due to the lack of information or the blocking and judicialization of privacy and security policies.

The perceptions presented were derived from the analysis of texts provided by the students and were further interpreted by the researchers. These interpretations were primarily informed by the researchers' extensive knowledge and experience within the realm of HDI. It's worth noting that additional reports obtained during the lecture activities also touched upon the same aspects. However, due to considerations of similarity and space constraints, these additional reports are not included here.

In a post-lecture activity, students were invited to use a mobile application built to assist in teaching activities. This application, proposed by Filgueiras *et al*. [2019], monitored the volume of data sent through the cell phone network interface and alerted the user every two megabytes of data were sent. The students used the application and reported the amount of data indicated by the application, which surprised them, as they were constantly warned by the application about the amount of transferred data.

Therefore, the initial conclusions of the application of the HDI themes in a lecture format were:

- They were efficient in arousing interest and reflection on the handling of personal data in software applications, since it was noted that the participants already had a certain knowledge on the subject, but did not care or did not consider its effects;
- Since the participants were undergraduate students in Computing subjects, it was possible to perceive a greater interest and concern with the fact that HDI is present in software projects and it was assumed that the participants will take this concern with them;
- A lecture is insufficient in time and content for in-depth discussion of issues related to the *design* project, however, it was assumed that with the knowledge of the need for HDI, the participants will look for strategies to support the inclusion of it in their design projects;
- The use of the data traffic monitoring application was of great help due to the fact that it allowed participants to have a perception of the volume of shared data.

In forthcoming activities, the plan is to customize and deliver analogous lectures for graduate courses within the field of computing. Additionally, these lectures will be extended to both undergraduate and graduate courses in interconnected domains like Law, Marketing, and Administration. These

[6] subscription of the students texts

disciplines also play a significant role in the utilization of personal data.

Next subsection shows the case study with discipline.

3.3.2 Content in the HCI undergraduation course

The second case study was conducted in the Computer Science and Information Systems undergraduation courses at the State University of Northern Paraná (UENP), in the academic year of 2021, under emergency remote teaching. In both courses, the disciplines were offered for senior students, with a workload of sixty hours in a semester. The courses were taught by the doctoral student, that was a professor at the university, and he was supervised by other professors pointed in the previous case study.

The HCI course is a mandatory subject of both courses and is allocated in the fourth period of the courses. Although the courses has a syllabus considering classic aspects of HCI, it can be adjusted to consider new contents, updates and adaptations, since HCI is a very versatile area and commonly receives updates from scientific research and Marketplace.

In the Computer Science course, activities were conducted with thirteen students and in the Information Systems course, with nine students, summing twenty-two. At the meeting, the age range of the students was 18 to 22 years old and considering the gender, the distribution was: Computer Science - 12 men and 1 woman; Information systems - 8 men and 1 woman. The students were not financially vulnerable and the Information Systems students already worked during the day, since the course was at night. Computer Science students, on the other hand, dedicated themselves to the course and to academic activities such as scientific initiation projects.

The contents proposed in Section 3.1 were presented and discussed with the students within the syllabus of HCI, always with the approach of prioritizing HCI content relating to HDI subjects. As a final activity, students were proposed to develop a document containing: (1) specification of requirements; (2) user profile identification (*persona* and empathy map); (3) task analysis; and (4) prototyping. This document should had included aspects of handling personal data when proposing an improvement in an existing tool or proposing a new software application solution.

Interesting and relevant proposals were presented such as: improvements in data use transparency; application to assist in the interpretation of policy and privacy texts; environments for sharing information about data privacy issues; and Tool for notifying violations in the handling of personal data. Deliveries included user's identification, with personas and empathy map, task analysis, with Hierarchical Task Analysis Models(HTA), pointed by Benyon [2011], and Solution prototyping.

In the requirements specification stage, students focused on the concerns of people, as data subjects, and on the obligations of data controllers. In all cases, the requirements were linked to the LGPD[7], papers and documents, such as: Article 48 (mandatory communication by the national authority about problems with data manipulation); and Article 9 (data subject's right to easy access to information on data handling). LGDP information are presented by Law [2020].

[7]https://www.gov.br/mds/pt-br/acesso-a-informacao/lgpd

Concerning the users identification, the creation of personas and empathy maps stood out. So, the students included concerns and characteristics related to privacy and the need to interact with the data and when the prototypes were on perspective, the focus was on the need for transparency. Students were engaged and concerned about the subject and provided options to assist data subjects in identifying personal data manipulation actions.

In the modeling and prototyping stages, students were encouraged to use HTA concepts and also design patterns for mobile devices in order to propose HCI solutions for their proposals. We chose to focus on mobile applications because the proposals were related to the use of personal data, that is more common with the use of cell phones, both for the production of data by data subjects and/or for the consumption of information. The results were considered satisfactory for the discipline.

At the end, a brainstorming activity was conducted with the students, they highlighted that: (1) they understood the concern with data privacy and with the actions of handling personal data in HCI projects; (2) HCI design and product adaptation needs are real, but they believe that the companies consider them due to regulations, but not due to user concerns; and (3) in the students opinions, it will not be simple to provide interfaces to support the aforementioned actions, given their particularities versus the needs of users. They also considered that this challenge can bring many benefits to HCI approaches, techniques and for the interfaces design such as: modeling techniques, interface components and evaluation techniques.

Thus, it was assumed that the inclusion and discussion of data privacy and HDI content in an HCI course was effective to abstract and practice design actions of privacy requirements and handling personal data in HCI projects. The inclusion of these contents did not harm the discussion of already consolidated contents of the discipline and allowed the students to conduct a relationship between the subjects and complement them efficiently.

For future works, with the full return of face-to-face classes, it is expected a more efficient discussion about this subject. It is also planned to update and revise the HDI contents and techniques that might interfere in HCI projects.

Next section presents the final considerations.

3.4 Considerations about HDI in the HCI course

The inclusion of HDI contents in HCI courses were considered satisfactory. As a course, it allowed a theoretical discussion and practical application of HDI concepts. Even with the focus on handling personal data, the students understood the need to be concerned with this topic in future projects and that concern for the individual that produces the data should be fundamental.

The lectures and workshops, although shorter, were efficient in presenting the subject to the participants and it is believed that it aroused the interest of some of them to carry out their activities in this subject. However, practical matters such as design, development and application of HDI con-

cepts could not be considered and analyzed due to time constraints.

Although it may arouse students' interest in this subject, dealing with HDI issues only as part of the HCI discipline fails to address several issues related to the handling of personal, corporate and scientific data, which will certainly be part of the graduates' professional activities. It is also necessary to consider that some subjects may not have such a strong relationship with HCI, such as the study of Statistics or Ethics, which may have a stronger relationship with other areas.

Therefore, it is safe to assure that there is a need to create a specific HDI course, which allows discussing the subject in Undergraduate and Postgraduate courses, at specialization and Masters/Doctoral levels, in a more in-depth and comprehensive way, with the discussion of various topics that support the performance of a data professional.

With the existence of this specific course, it is assumed that an appropriate preparation of a data professional occurs, since that these subjects need to be considered as significant when being analyzed with the related contents of other courses such as Software Engineering, Database, Administration and HCI itself. This can enable the student to be able to act consistently, safe and effective in designing or implementing data systems.

Next section presents a proposal for a minimum curriculum for teaching Human-Data Interaction (HDI).

4 Human-Data Interaction Course

This section presents a proposal for the Human-Data Interaction (HDI) discipline for undergraduate and for stricto or lato sensu postgraduate courses in the area of Computer Science, as well as suggestions for lectures and dissemination workshops.

The HDI discipline aims to immerse students in conceptual, human, scientific, methodological and technical aspects, that must be presented according to the level at which the discipline is offered, seeking to support the students´ academic and professional activities, aiming to present contexts that make their participation more efficient, ethical and innovative.

Considering that the course aims to approach different levels of Computing courses, the Number of Hours (NH) of course proposed for each approach, as well as the distribution of themes, will be adjusted considering the practice in courses in which the experiments presented in Section 3 were conducted. Although this paper presents a NH proposal for each topic, it is totally flexible and understandable that possibles adjustments on the hours allocated for each content, as well as for the discipline, can be defined. This is related to the particularities that each discipline or educational institution may have.

The HDI course was organized into four Groups and nine themes. The groups are (1) Fundamentals of HDI; (2) HDI applications; (3) HDI Systems, Methods and Principles; and (4) HDI Research, Innovation and Education. Each of the groups and their respective topics will be discussed in the next subsections. The themes of each group will also be further discussed.

4.1 Group 01 - Fundamentals of HDI

The first group refers to the foundations of the HDI and aims to present and contextualize the data use ecosystems and the relationships established with human, social, technological and corporate issues. In this group, two topics should be discussed:

4.1.1 Theme 01: HDI Contextualization

In this theme, the presentation and contextualization of HDI should occur for students in order to instigate them about the challenges and opportunities in data manipulation environments. This theme should include, at a minimum, the following topics:

- **Relationship between humans, data and Artificial Intelligence (AI)** - the manipulation of data with the help of AI and the several possibilities of information, products and services generated from them and the impact they can have on people, who position themselves as data producers or consumers. These data can be also used as inputs for the training and execution of AI models to generate information about people, processes and organizations;
- **Origin of data** - present the vast possibility of data sources such as personal data, IoT, Smart Things, health data, corporate data, Industry 4.0, Government and others;
- **History** - discuss historical aspects of the data usage and approach the key issues that led to the growth of interest in data. It is recommended to discuss historical issues of Big Data, Infovis and the interest in Personal Data;
- **Demand for data scientists and challenges in the industry** - In order to provide a practical view of the issues discussed, recent reports might be presented to allow understanding the need of and the expected profile of scientists and other professionals in the area.

It is also important to present the necessary prerequisites for the course, since HDI inherits aspects from other areas and commonly requires adjustments in various methods and processes normally applied to conventional projects, such as:

- Software Engineering - requirements, modeling and testing, to deal with issues related to HDI projects;
- Design Process - Usability and User Experience, focusing on the design of user interfaces and experiences that allow interaction with tools and/or information;
- Relational and NoSQL Database, to understand how the data will be organized and how it can be considered in an HDI project.

This theme does not need to be extensive, as the subjects will be discussed in greater depth in the following steps. Thus, it is proposed that it can be presented in 04 hours for undergraduate or post stricto sensu courses and 03 hours for lato sensu courses.

4.1.2 Theme 02 - Aspects of perception and cognition for HDI

This theme aims to expand students' knowledge of the human perceptual and cognitive processes that underpin data visualization and decision-making design heuristics, so that students understand the impact of design choices on the appreciation and use of visual interfaces.

Topics for this theme could be:

- **Human vision and pre-attentive attributes** - notions of visual perception and pre-attentive processing of visual attributes such as color, contrast, shape and position are discussed;
- **Gestalt principles** - the principles of emergence, invariance, figure/ground and multistability are presented, as well as their relationships with aspects of grouping, filling and continuity.
- **Attention** - the attention mechanism, its limitations and ways to capture attention by design are discussed.
- **Decision-making** - the mechanisms that define how people make decisions are discussed. Aspects of biases (confirmation bias, availability bias) and constraints such as stress and information quality are also presented. Types of decision-making: rational, intuitive and heuristic; and the role of data in these processes are also discussed.

This theme can be taught in 4 hours for undergraduate courses, 3 hours for lato sensu postgraduate courses, and should be extended to 8 hours with research activity in stricto sensu postgraduate courses.

4.2 Group 02 - HDI Applications

This group aims to discuss how HDI is inserted in different data manipulation approaches such as personal data, data embedded in 3D interactions and information visualization. HDI is also discussed in projects such as biodiversity, agriculture, health monitoring, Internet of Things, among others. This topic is based on the need to understand the possibilities of data manipulation from a practical point of view and on the considerations about people's relationships and aspects related to these actions and future perspectives. This group was divided into two themes, presented in the next subsections.

4.2.1 Theme 03 - Handling of Personal Data

The purpose of this theme is to discuss aspects related to the handling of personal data. Personal data is a record that identifies or allows identifying a person, his behavior or his characteristics (Mortier *et al*., 2016). The following topics can be discussed under this theme:

- **Concepts and origin of personal data** - this topic discusses how technological evolution has led to a great immersion of people in resources such as Smartphone, Smart Home, Internet, among others. It should address how people became dependent on the technology and its use it to perform everyday activities such as entertainment, household activities, teaching activities, professional activities, among others.
- **Manipulation actions** - Manipulation is considered as any action performed on personal data such as: collection/acquisition, processing, sharing, disclosure, storage and disposal (Toledo, 2020). Thus, this topic seeks to present the various manipulation possibilities with examples of different possible domains, among them:
  - **Quantified Self**: collection of an individual's personal data, processing and production of information about the individual him/her self for his/her use, such as monitoring health parameters;
  - **Interpersonal data manipulation**: companies that collect personal data, produce information and make it available to other people with common interests or as a way to establish a relationship between these people;
  - **Commercial Manipulation**: companies collect and manipulate personal data in order to serve their commercial and financial interests, such as consumer profiling and targeting products and services.
- **Regulations** - this topic aims to encourage a discussion about the European General Data Protection Regulation (GDPR)and the brazilian *Lei Geral da Proteção de Dados* (LGPD), among other regulations of interest. In particular, points related to data system requirements for compliance with regulations, as well as issues directly related to individuals such as Transparency, Privacy, Action and Negotiation must be extensively discussed;
- **Personal Data Transparency** - in this topic it should be discussed the aspects related to how information about the handling of personal data is made available to their holders. The concept of Transparency of Personal Data must be clarified, as well as its operational challenges, such as improving perception capacity, the use of graphical interfaces, icons and new models of interaction with the user. The implications and impacts of Transparency for controllers and individuals must also be addressed.
- **Action and Negotiation** - in this topic, aspects related to the Negotiation Action should be addressed since, considering that the individual has clear and perceivable information about the manipulation of his/her personal data, he/she has grounds to act in order to negotiate, restrict or cancel any permissions and contracts for handling their data.

This theme has a strong theoretical appeal, but with possibilities for practical activities, which may include:

- Conducting analyzes and evaluations of Privacy and Security Policies and technological resources for the Action and Negotiation of existing applications;
- Studies of tools, applications and environments that manipulate personal data, in order to identify their objectives, ethical and technological issues of the process.

It can be taught in 8 hours for undergraduate courses, 3 hours for lato sensu postgraduate courses and should be ex-

tended to 16 hours with research activity in stricto sensu postgraduate courses.

4.2.2 Theme 04 - Information Visualization

This theme is extensive and should be presented in some other course or even program for students interested in HDI. Considering the introduction to Infovis in a single course, the essential topics are listed.

- **Audience and View Types** - This topic discusses different view types depending on the intent of the audience. Two types stand out - visualization of an exploratory nature, whose audience is generally formed by data scientists and analysts, who want to interact with the data and with different forms of visualization in order to extract insights and new knowledge. The second type is visualization of a narrative nature, or infographics, whose audience is generally formed by consumers of information, who need to understand facts from data. In these cases, the narrative involves highlighting data and highlighting relationships.
- **Infovis design** - This topic is quite extensive and should contain at least the following concepts:
  - Definition of categorical and numerical variables, with examples;
  - Representation of variables and visualization grammars, presenting the most common types of graphs and their use;
  - Presentation of variables, with examples of dashboards;
  - Interaction capabilities and visual data analysis
  - Data visualization tools
  - Examples of visual languages in different application areas, for example in epidemiology, cartography and meteorology.
- **Design systems** - This topic gathers information on how to accelerate the visualization design process in an enterprise by creating design systems containing visual identity, design patterns for graphics and dashboards; visualization of missing data and uncertainties;
- **Infovis quality** - In this topic, aspects of visualization quality are discussed, which include aspects of effectiveness, intelligibility, visual integrity and aesthetic superiority.
- **Narrative visualizations** - In this topic, examples and techniques are presented for the construction of information presentations with the purpose of narrative, such as data journalism and infographics.
- **Accessibility** - In this topic, considerations are presented about the accessibility related to graphs and tables of data, as well as interactive data manipulation tools.

This topic, due to its length and relevance, should occupy 12 hours in undergraduate courses, 09 hours for and post-lato sensu courses, and should be extended to 24 hours with research activity in stricto sensu post-graduate courses.

4.3 Group 03 - HDI System, Methods and Principles

This group aims to present tools and applications used and applied in HDI. By the term **systems** it is understood as software and hardware applications used by professionals, which enable actions from the planning stage, collection to the dissemination and sharing of information. The set of systems currently available for the visual analysis of data and visualization design, commercial and open source, allow the construction of extremely rich, productive and interactive visualizations, which can provide great insights for companies and people.

As for the term **applications**, one should consider software and hardware products and services that can be used by people for personal and professional activities. In this line, changes and adaptations in traditional methods and techniques of Software Engineering, Database and Human-Computer Interaction are necessary, since HDI application projects have particularities such as: technologies, legislation, business, processes and people. These particularities must be understood and abstracted in order to support the effective construction of a data product or service.

4.3.1 Theme 05 - Systems

In this group should be presented the HDI systems, with a practical nature, for the development of ”know-how”. Thus, it is assumed that the classes are taught in a laboratory, using computers and some software. The use of free tools is encouraged, as well as the establishment of educational agreements with companies that could provide commercial data visualization tools. The objective of this laboratory is to allow the students to practice the design principles presented in the course, producing visualizations from datasets, evaluating the quality of their results and exploring the interaction capacity of the tools. The following topics might be foreseen in this theme.

- **Classic views** - use of tools to produce classic graphics and essential operations from a dataset.
- **Construction of dashboards** - use of tools to, based on a dataset, produce a panel with diverse and connected graphs, aiming to address a specific problem as an example of decision making.

This theme can be taught in 8 hours for undergraduate courses, 3 hours for lato sensu postgraduate courses, and should be extended to 16 hours with research activity in stricto sensu postgraduate courses.

4.3.2 Theme 06 - HDI Methods and Principles

This theme seeks to promote discussions on approaches and principles to support HDI development projects and people’s appropriate use of data.

In the literature it is possible to identify papers and articles that discuss specificities and necessary improvements in Software Engineering approaches, HCI, Database for data-oriented applications. Related to this subject, for this paper,

Gorton *et al*. [2016] and Ji *et al*. [2020] can be cited. With regard to aspects of people's privacy in projects, consolidated contributions with a strong application bias in HDI can be observed, such as: principles as Privacy by Design; and the Privacy User Experience.

For this topic, the following topics should be addressed:

- **Software Engineering for HDI** - in this topic, the basic Software Engineering concepts should be discussed to analyze how they are, or at least could be, applied to the particularities of HDI projects. It is recommended to reflect on the evolution of HDI and the consequent need for improvements in the processes of building applications for HDI. At a minimum, the following should be addressed:
  - Requirements Engineering Issues for HDI;
  - Software Architecture for HDI;
  - Tests and Software Quality for HDI.
- **Privacy by Design (PbD)** - this topic should discuss the principles of PbD so that the product or service design can include user privacy as a mandatory requirement;
- **Privacy User Experience** - discuss elements for creating user interfaces and experiences that can support individuals in understanding and managing the handling of their personal data and the consequent control of their privacy, security and freedom.
- **Data Literacy** - discuss aspects related to training people to handle with data, both in the sense professionals in companies and overall people who produce and use data through applications or other commonly used computing environments. The discussion of this content can be based on the principles of Data Literacy proposed by authors cited in this work.

This topic must be covered in 12 hours for undergraduate courses, 3 hours in postgraduate courses in the broad sense, and must be extended to 20 hours with research activity in stricto sensu postgraduate courses.

4.4 Group 04 - Research, Education and Innovation in HDI

The HDI area is a booming one, as it accompanies the accelerated growth of data manipulation in addition to people's interest and need for data, information and knowledge. Supported by this need, it is very important to discuss the perspectives, challenges, evolutions and contributions in methods, techniques, tools and strategies of the HDI.

The activities of this group can be of a practical nature and include specific ones that can be conducted in the classroom: exercises, seminars and Final Course project. Related to these activities, the students can analyze and discuss scientific papers, reports, innovation projects and/or any other artifacts that may guide HDI's future. In addition to analyzing the existing documents, academics can work to develop and present proposals for new approaches, tools, methods, techniques, products or services aimed at HDI.

The carrying out of projects may also be considered since, in a theoretical approach, it addresses the future of HDI and the aspects that surround it. The following topics for study and discussion are suggested:

- **Research**: analyze studies related to the scientific evolution of HDI concepts, methods, techniques, tools and applications. This topic can be discussed more carefully in undergraduate and stricto sensu postgraduate courses, since scientific research is an inherent part of their curriculum. It is also recommended to carry out activities in which students can propose research projects within the area of HDI. Examples of discussion can be: (1) methods for immersing users in data analysis; and (2) analyze the adequacy of Software Engineering and Human-Computer Interaction processes for HDI;
- **Innovation**: discuss and propose, even superficially, innovative projects for HDI, especially those that can bring about changes in the way data is handled and, consequently, in the way people are impacted by data and information . Considering that it has become more common in Higher Education Institutions to have Technological Innovation Agencies, a working partnership can be productive. The discussion of innovation can be very productive for lato sensu courses, since the participants tend to be more active in the general market/companies when compared to the strictu sensu ones, that are more related to the academia;
- **Teaching**: address strategies to teach people (regular citizens and data professionals) to interact and use data systems effective and efficiently, so that they can contribute to their daily activities.

With regard to the NH (Number of Hours of classes) of this group, for undergraduate courses, which have a fundamental training profile, a 12-hour is recommended; for stricto sensu postgraduate courses, which are based on the training profile of researchers, it is recommended 32 hours. For lato-sensu courses, which have an aspect of practical and immediate application, but with a more restricted NH, a 3-hour is recommended. In all course profiles, but in particular the lato-sensu courses, due to the time restriction, this group can be developed partly in the classroom, partly in the discipline's activities, final or extension projects.

4.5 Summary and Totalization of Hours

The NH proposal presented in the previous sections were selected considering undergraduate and graduate courses discussed in Section 3, namely:

- **Undergraduate course**: 60 hours;
- **Professional graduate course (*lato sensu* program)**: 30 hours.
- **Research-base graduate course (*stricto sensu* program)**: 120 hours;

The NH proposed for each theme was presented in the corresponding subsections. For better presentation purposes, they are also informed in Table 1. For purposes of visual organization, the abbreviations UC will be used, for Undergraduate course; RBC for Research-base graduate course; and PGC for Professional graduate course.

Table 1. Carga Horária proposta para HDI de acordo com o tipo do curso

Theme	UC	PGC	RBC
HDI Contextualization	4	3	4
Aspects of perception and cognition for HDI	4	3	8
Personal data manipulation	8	3	16
Information Visualization	12	9	24
HDI Systems	8	6	16
HDI Development and Evaluation Methods	12	3	20
Research, Inovation and Education	12	3	32
Total	**60**	**30**	**120**

Next section presents the recommended bibliographical references for the discipline of HDI.

4.6 Bibliographic References for HDI discipline

As mentioned before, HDI is a relatively new field and inherits characteristics from areas such as HCI, Data Science, Psychology and Sociology. So, specific bibliographies, especially books, are still rare. For scientific papers, it is already usual to find studies involving the term HDI more specifically.

The indicated books refer to the related areas, which influenced the definition of HDI. Scientific papers that present the state of the art and proposed a series of significant contributions for HDI are also indicated. It is known that HDI is evolving as the use of data also also evolves and grows and it is known and acceptable that professors insert other bibliographies to meet their particularities in their classes. It is also known that new materials will be published and may integrate the final bibliography that can be proposed.

The minimum bibliography suggested for HDI course is shown in Table 2 and includes books from areas related to HDI, Organizational Reports on actions with data and papers that, when we were producing this paper, were still based on their authors' previous experience with teaching HDI. Anyway, they can provide significant contributions to the understanding and application of HDI.

Next subsection presents about the adequacy of the HDI course with Brazilian Computing Society and Brazilian Education Council.

4.7 Adequacy of the Course with the Brazilian Computing Society and Brazilian Education Council

As this paper presents a HDI course proposal, for achieving this goal, not just the presentation of the aspects to be discussed in this course, but also the considerations related to the technical documents, presented in curricular guidelines for graduate and undergraduate courses at Brazilian Computing Society (SBC), acronym for the Portuguese (*Sociedade Brasileira de Computação)* was considered. This SBC´s document also refers to the training references for the courses in computing areas, such as undergraduate and postgraduate degrees in Computer Science, Computer Engineering and Information Systems.

So, when the Ministério da Educação Superior - Conselho Nacional de Educação [2016], in its Resolution nº 5, of November 16, 2016 proposed by SBC, was taken into account, the items presented at the Article 4o were highly considered for the HDI course proposal. The topics that represent this article are:

1. Knowledge of social, professional, legal, ethical, political and humanistic issues;
2. Understanding the impact of computing and its technologies on society with regard to meeting and strategically anticipating society's needs;
3. Critical and creative vision in identifying and solving problems contributing to the development of their area;
4. The ability to act in an entrepreneurial, comprehensive and cooperative manner in meeting the social demands of the region where it operates, Brazil and the world;
5. To rationally use the available resources in a transdisciplinary way;
6. Understanding the needs for continuous updating and improvement of its skills and abilities;
7. The ability to recognize the importance of computational thinking in everyday life, as well as its application in other domains and be able to apply it in appropriate circumstances; and
8. The ability to act in a globalized world of work.

The aspects mentioned in this Article 4 point out that Computing professionals should not limit themselves to the technical aspects, but should also consider how human and social factors cross-cut in the development, usage and maintenance of computational solutions. In this context, the proposal of the HDI course aims to enable the student to apply knowledge that goes beyond the creation of models, algorithms, or databases, but includes the understanding of human, social, ethical, and political aspects that data manipulation can provide for individuals and society.

The HDI course collectively seeks to cultivate analytical, critical, and creative proficiencies for constructing data-driven computational solutions that cater to both business requirements and societal imperatives. This is achieved through a conscientious and ethical utilization of data, with a constant emphasis on placing individuals at the core of the data ecosystem. Furthermore, the envisioned approach in HDI is designed to empower students in recognizing and incorporating diverse everyday contexts into a data philosophy that informs pivotal decisions on a global level.

The proposed HDI course is intended to attract professionals from both undergraduate (Bachelor's and Licentiate) and postgraduate (stricto or lato sensu) programs in fields such as Computer Science, Computer Engineering, and Information Systems.

As outlined in the Training References for Undergraduate Computing Courses, outlined by Zorzo *et al*. [2017], professionals within these domains are expected to cultivate proficiency in systems development and IT management. This entails devising solutions that align with organizational strategies, overseeing Information Systems (IS) projects, and adhering to optimal management and governance practices for

Table 2. Suggested Bibliography for the HDI Course.

Type	Referenece
Books	Gomes, Elisabeth; BRAGA, Fabiane. Inteligencia Competitiva Tempos Big Data. Editora Alta Books, 2017.
	Amaral, Fernando. Introdução à Ciência dos Dados. Rio de Janeiro: Alta Books, 2017.
	Clegg, Brian. Big Data: How the Information Revolution Is Transforming Our Lives. Icon Books, 2017.
	Knaflic, C. N. Storytelling com dados: Um guia sobre visualização de dados para profissionais de negócios (2nd ed.). Alta Books, 2019
	Cairo, A. Functional Art, The: An introduction to information graphics and visualization. Analytics Press, 2013.
	Morrow, J. Be Data Literate: The Data Literacy Skills Everyone Needs to Succeed. Kogan Page, 2021.
Chapter	Mortier, R., Haddadi, H., Henderson, T., Mcauley, D., Crowcroft, J., & Crabtree, A.. Human-Data Interaction: The Encyclopedia of Human-Computer Interaction. The Encyclopedia of Human-Computer Interaction, 1–48. Retrieved from https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed, 2016
Reports	Accenture. The Power of the Data-Driven, 2019.
	Accenture. Multi-speed data and analytics, 2021.
	Accenture. Big Success With Big Data, 2014.
	Cavoukian, A., Privac y by Design. Identity in the Information Society, 3(2), 1–12, 2010
Papers	Mortier, R., Haddadi, H., Henderson, T., McAuley, D., & Crowcroft, J., Human-Data Interaction: The Human Face of the Data-Driven Society, 2015, https://doi.org/10.2139/ssrn.2508051
	Hornung, H., Pereira, R., Baranauskas, M. C. C., & Liu, K., Challenges for Human-Data Interaction – A Semiotic Perspective BT - Human-Computer Interaction: Design and Evaluation. Human-Computer Interaction: Design and Evaluation, 1, 37–48, 2015, https://doi.org/10.1007/978-3-319-20901-2
	Amo-filva, D., García-peñalvo, F. J., & Chen, J., Towards an ethical data literacy proficiency : a Moodle logs analytical tool. Proceedings of XII International Conference on Virtual Campus (JICV), (1), 2022–2024, 2022.
	Barth, P. (n.d.). The Seven Principles of Data Literacy The Seven Principles of Data Literacy A Blueprint to accelerate your business toward its data-driven future Foreword from Qlik. Accenture.
	Murmann, P., & Fischer-Hübner, S., Tools for Achieving Usable Ex Post Transparency: A Survey. IEEE Access, 5, 22965–22991, 2017, https://doi.org/10.1109/ACCESS.2017.2765539
	Coleti, T. A., Corrêa, P. L. P., Filgueiras, L. V. L., & Morandini, M., TR-Model. A Metadata Profile Application for Personal Data Transparency. IEEE Access, 8(1), 75184–75209, 2020, https://doi.org/10.1109/ACCESS.2020.2988566

these systems. An integral facet of these solutions involves the creation of IS software, necessitating skills in specifying, producing, and implementing such solutions.

Graduates can leverage their expertise in data and information management to provide valuable support to organizations in various capacities. This includes aiding in the selection of appropriate technologies, specification processes, as well as the implementation and utilization of databases.

Therefore, this course is designed to equip students with the skills required to engage in data-focused projects across various stages, including problem identification, solution development, ideation, construction, evaluation, and project utilization, among others. Furthermore, the HDI course holds versatility in its applicability across diverse undergraduate programs, underlining the imperative of handling data transparently, appropriately, and ethically to fulfill distinct data-related demands aligned with the respective course profiles.

For instance, in a Computer Science program, the course might delve into strategies for creating computational tools tailored to HDI, whereas in a Licentiate in Computing program, the focus could be on executing data projects with educational objectives.

This HDI course proposal was crafted with the intention of fulfilling the aforementioned descriptions. Its primary objective is to equip participants with a comprehensive understanding of Data Management principles. This encompasses the imperative of collecting, storing, presenting, and managing data in a transparent and effective manner. Throughout the course, there is a consistent emphasis on recognizing the far-reaching effects of data utilization on both the environment and individuals in their day-to-day lives.

Next, we present the considerations about the HDI disciplines.

4.8 Considerations about the HDI discipline

Section 4 presented a proposal for a HDI discipline for undergraduate, lato sensu postgraduate and strictu sensu postgraduate courses. The subject's proposal is in line with the wide use of data that has become the basis for the operation of several computer systems, and the greater reach and interference of data manipulation in people's lives, as producers or consumers.

With a specific HDI discipline, it is expected to contribute to the formation and development of graduates with knowledge in data projects, increase their relationships and impacts

with people, so as not to focus only on architectural aspects[8] such as Databases, Algorithms and Machine Learning. It also aims to observe how these projects are used or can influence people's lives.

The proposed content will allow professors to discuss content related to developers and people, who as producers and consumers of data and information end up positioning themselves as the core of HDI projects. In this sense, the proposal aims to promote to graduates that HDI is not, solely, the Data Science discipline, just as it is not the HCI discipline, but an intersection of them with their particularities that must be observed by data professionals.

Finally, given the relevance of the topic within the area of Computing, it is expected that, in a short term, the discipline of HDI may be included in the curricula of undergraduate Computing courses, lato and strictu sensu postgraduate. These last two are easier to implement in relation to graduation due to bureaucratic issues of curriculum adequacy.

Next section presents some proposals for HDI lectures and workshops.

5 HDI as workshop and lecture

For situations in which time and resources are not available for a course, such as events, workshops, or isolated classes, the discussion on HDI can also be done through lectures and workshops. These both aim to bring an overview and/or a practical activity on the topic of HDI, to guide and/or encourage your interest and immersion in the subject.

For a Lecture proposal, the content presented may cover theoretical and fundamental concepts in the subject, with examples of HDI applications. This aims to show the relationship between data use and corporate and personal aspects. It is suggested as contents for the lectures:

- HDI context - main concepts, data, human and technological aspects;
- HDI application - examples of projects, use of Machine Learning;
- Personal data - transparency, action and negotiation in the handling of personal data;
- Privacy by Design and Privacy User Experience;
- Data Literacy

A maximum period of 2 hours of lecture is suggested, given its theoretical profile and respect for the participants' ability to concentrate.

The **workshops** must have a practical aspect, with the support of some computational tool such as Public-Tableau [9], in which participants can practice a process of acquiring, transforming and/or presenting information to users. Considering these characteristics, the following activities are suggested:

- Contextualization of HDI from the point of view of data collection and delivery of information to users;
- Select a case study for demonstration;
- Get a set of data and import it into the Public-Tableau tool[10];
- Apply Infovis concepts to build visuals to inform users.

A workshop lasting between 2 and 4 hours is recommended and the deepening of the mentioned topics can be adjusted according to the availability of time and resources. However, it is strongly recommended to focus on practice rather than theoretical concepts, in addition to using the participants' computer lab or computers.

If time and resources are available, the combination of the lecture and the workshop may be even more efficient in contextualizing and exemplifying HDI projects, providing more concrete examples that can be discussed in depth in order to support the acquisition of the basic knowledge necessary to future activities by students.

Next section presents the conclusions and expectations of this paper.

6 Conclusions, Expectations and Future Work

This paper introduces a structure for a course on Human-Data Interaction (HDI) course, a subject that has emerged in response to the exponential proliferation of data-driven projects. In an integrated manner, this subject has permeated people's daily activities, significantly impacting their decisions and experiences through the use of products and access of services reliant on data utilization. This study not only proposes a comprehensive curriculum for the HDI course but also offers streamlined alternatives for instances where a complete course implementation might not be feasible.

Furthermore, this work is an extension of the paper titled **"Insertion of Human-Data Interaction Contents in the Human-Computer Interaction (HCI) Course"** by Coleti *et al*. [2022], originally published in the esteemed Annals of the XXX Workshop on Computing Education.

We held the conviction that the manipulation of both corporate and personal data ought to become increasingly thorough and extensive. This heightened manipulation is anticipated to yield a proliferation of digital products and services seamlessly integrated into the everyday routines of both individuals and domain experts. Consequently, these innovations stand to exert a substantial impact on routine tasks. Envisioning this trajectory, it becomes evident that numerous sectors possess the potential to enhance their offerings significantly through the integration of data-driven applications into their respective domains.

This evolution, however, mandates the presence of capable professionals well-versed in comprehending and applying computational methodologies and techniques. Furthermore, these professionals must be attuned to the intricate interplay of human, ethical, and social dimensions inherent in the realm of data.

Through the training of HDI designers, we envisaged the creation of computational tools accessible to a diverse spectrum of users, irrespective of their expertise levels. These tools, facilitated by an effective data interaction strategy, enable users to execute their tasks efficiently while upholding privacy and security protocols.

[8] Aspects focused on a discipline of Data Science.

[9] https://public.tableau.com/app/discover

[10] Or another one of your preference

The proposed course design can be seamlessly incorporated into both undergraduate and postgraduate programs, encompassing both ”lato sensu” and ”stricto sensu” categories. This integration could equip students with the necessary skills to engage in data-centric projects effectively. The distinctive feature of the HDI curriculum, as opposed to both Data Science and HCI courses, is its emphasis on amalgamating aspects intrinsic to both fields. This fusion centers on the end user, highlighting their role in both producing and consuming data and information.

This course endeavors to nurture students with a profound understanding of human-centric factors within data projects. It aims to instill an awareness that algorithms, database structures, and visualization models must be meticulously tailored to suit their intended audience and the contextual framework of their application. Through this interdisciplinary approach, graduates will emerge equipped to holistically integrate human considerations into every facet of data-related endeavors. It is also expected that the introduction of this discipline will serve as a mechanism to encourage undergraduate and postgraduate programs to incorporate more content related to data projects that emphasize human aspects. This inclusion aims to inspire educational institutions to integrate considerations of the human element into their endeavors involving data-driven projects.

We had anticipated that this discipline will serve as a catalyst for the inclusion of data studies within the foundational stages of education, promoting data literacy. Recognizing that today's youth are avid technology users, there is a pressing need to expose them to suitable applications and cultivate a nuanced understanding of utilizing data responsibly.

Envisioning the road ahead, a prospective avenue of this project entails the development of a Data Literacy course tailored for primary and secondary education levels. This initiative not only aims to gauge the aspirations and requirements of young learners but also strives to furnish them with comprehensive guidance concerning data utilization. By engaging with data in a constructive and informed manner, these students will be better equipped to navigate the contemporary digital landscape, thereby fostering a generation of digitally savvy and ethically conscious individuals.

Finally, we wish to emphasize that the course approach being presented is not focused on competences, as the SBC proposal is systematized. In this sense, we also want to highlight that this competency-based analysis should be revisited in light of both the Association for Computing Machinery [2020] and ACM Data Science Task Force [2021], and this might be one of our future activities to better improve this HDI course proposal. Particularly, we intend to address the topics related to Knowledge Area Analysis and Presentation (Knowledge, Skills and Dispositions) that are presented in its Appendix A.

References

Accenture (2014). Big Success With Big Data. pages 1–12.

Accenture (2019). The Power of the Data-Driven.

Accenture (2021). Multi-speed data and analytics. *Accenture*. DOI: 10.1201/9781003291800-9.

ACM Data Science Task Force (2021). *Computing Competencies for Undergraduate Data Science Curricula*. Association for Computing Machinery, New York, NY, USA.

Allard, S. (2012). DataONE: Facilitating eScience through Collaboration. *Journal of eScience Librarianship*, 1(1):4–17. DOI: 10.7191/jeslib.2012.1004.

Amaral, F. (2016). *Introdução à Ciência dos Dados*. Alta Books, Rio de Janeiro.

Amo-filva, D., García-peñalvo, F. J., and Chen, J. (2022). Towards an ethical data literacy proficiency : a Moodle logs analytical tool. *Proceedings of XII International Conference on Virtual Campus (JICV)*, (1):2022–2024.

Arass, M. E. and Souissi, N. (2018). Data Lifecycle : From Big Data to Smart Data Data Lifecycle : From Big Data to Smart Data. *Proceedings of 5TH Edition International IEEE Congress on Information Science and Technology (CiSt'18)*, (November):80–87.

Association for Computing Machinery (2020). Computing Curricula 2020. Technical report.

Baig, A. (2020). What is Privacy UX How to Implement Privacy-Aware Design Framework? *AltexSoft - Software and Engineering*.

Barth, P. (2022). The Seven Principles of Data Literacy The Seven Principles of Data Literacy A Blueprint to accelerate your business toward its data-driven future Foreword from Qlik. *Accenture*.

Basarudin, N. A., Yeon, A. L., Yusoff, Z. M., Dahlan, N. H. M., and Author, N. M. (2017). Smart home users' information in cloud system: A comparison between Malaysian personal data protection act 2010 and EU general data protection regulation. *Malaysian Construction Research Journal*, 2(2):209–222.

Beber, M. A., Ferrero, C. A., Fileto, R., and Bogorny, V. (2017). Individual and Group Activity Recognition in Moving Object Trajectories. *Journal of Information and Data Management*, 8(1):50. DOI: 10.5753/jidm.2017.1606.

Bellamy, B. and Alonso, C. (2016). Reframing data transparency _. *Centre for Information Policy Leadership and Telefónica Senior Roundtable*, 1(June):1–20.

Benyon, D. (2011). *Interação Humano Computador*. Pearson Education, São Paulo.

Boscarioli, C., Silveira, M. S., Prates, R. O., Bim, S. A., Diniz, S., and Barbosa, J. (2012). Currículos de IHC no Brasil : Panorama Atual e Perspectivas. *... em Computação, ...*, (2007):1294–1303.

Cafaro, F. (2012). Using embodied allegories to design gesture suites for human-data interaction. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, page 560. DOI: 10.1145/2370216.2370309.

Cairo, A. (2013). *Functional Art, The: An introduction to information graphics and visualization*. Analytics Press.

Cavoukian, A. (2020). Understanding how to implement privacy by design, one step at a time. *IEEE Consumer Electronics Magazine*, 9(2):78–82. DOI: 10.1109/MCE.2019.2953739.

Cham, H., Malek, S., Milow, P., and Song, C. (2022). Developing an ecological visualization system for biodiversity data. *All Life*, 15(1):500–511. DOI:

10.1080/26895293.2022.2066195.

Choe, E. K., Lee, N. B., Lee, B., Pratt, W., and Kientz, J. A. (2014). Understanding Quantified-Selfers' Practices in Collecting and Exploring Personal Data. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, pages 1143–1152. DOI: 10.1145/2556288.2557372.

Clegg, B. (2017). *Big Data: How the Information Revolution Is Transforming Our Lives*. Icon Books.

Coleti, T., Morandini, M., and Filgueiras, L. (2022). Inserção de conteúdos de interação humano-dados e privacidade de dados na disciplina de interação humano-computador. In *Anais do XXX Workshop sobre Educação em Computação*, pages 181–191, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/wei.2022.223236.

Coleti, T. A., Corrêa, P. L. P., Filgueiras, L. V. L., and Morandini, M. (2020). TR-Model. A Metadata Profile Application for Personal Data Transparency. *IEEE Access*, 8(1):75184–75209. DOI: 10.1109/ACCESS.2020.2988566.

Columbia University Coms (2022). Systems for human-data interaction.

Cradock, E., Stalla-Bourdillon, S., and Millard, D. (2017). Nobody puts data in a corner? Why a new approach to categorising personal data is required for the obligation to inform. *Computer Law and Security Review*, 33(2):142–158. DOI: 10.1016/j.clsr.2016.11.005.

Cybis, W. d. A., Holts, A. B., and Faust, R. (2015). *Ergonomia e Usabilidade: Conhecimentos, Métodos e Aplicações*. Novatec Editora, São Paulo.

Efroni, Z., Metzger, J., Mischau, L., and Schirmbeck, M. (2019). Privacy icons: A risk-based approach to visualisation of data processing. *European Data Protection Law Review*, 5(3):352–366. DOI: 10.21552/edpl/2019/3/9.

Elmqvist, N. (2011). Embodied Human-Data Interaction. In *CHI 2011 Extended Abstracts on Human Factors in Computing Systems*, pages 1–4.

Evequoz, F. and Lalanne, D. (2006). Personal Information Management through Interactive Visualizations. *Proceedings of Doctoral Colloquium of IEEE Information Visualization Conference (Infovis-DC 2007)*, (August):158–160.

Few, S. (2016). Data Visualization for Human Perception. *The Encyclopedia of Human-Computer Interaction*, 2.

Filgueiras, L. V. L., Leal, A. S. F., Coleti, T. A., Morandini, M., Correa, P. L., and Alves-Souza, S. N. (2019). Keep System Status Visible: Impact of Notifications on the Perception of Personal Data Transparency. *Human-Computer Interaction. Perspectives on Design*, 1:513–530.

Gomes, E. and Braga, F. (2017). *Inteligencia Competitiva em Tempos Big Data*. Alta Books.

Gorton, I., Bener, A. B., and Mockus, A. (2016). Software engineering for big data systems. *IEEE Software*, 33(2):32–35. DOI: 10.1109/MS.2016.47.

Haddadi, H., Chaudhry, A., Crowcrof, J., Howard, H., Mortier, R., and Mcauley, D. (2015). Personal Data: Thinking Inside the Box. *Aarhus Series on Human Centered Computing*, page 8. DOI: 10.7146/aahcc.v1i1.21312.

Hartson, R. and Pyla, P. (2012). *UX Book. Process and guidelines for ensuring a quality user experience*. Morgan Kaufmann - Elsevier.

Holtz, L. E., Nocun, K., and Hansen, M. (2011). Towards displaying privacy information with icons. *IFIP Advances in Information and Communication Technology*, 352 AICT:338–348. DOI: 10.1007/978-3-642-20769-3$_2$7.

Hornung, H., Pereira, R., Baranauskas, M. C. C., and Liu, K. (2015). Challenges for Human-Data Interaction – A Semiotic Perspective BT - Human-Computer Interaction: Design and Evaluation. *Human-Computer Interaction: Design and Evaluation*, 1:37–48. DOI: 10.1007/978-3-319-20901-2.

Hosseini, M., Shahri, A., Phalp, K., and Ali, R. (2016). Foundations for transparency requirements engineering. In Daneva, M. and Pastor, O., editors, *Requirements Engineering: Foundation for Software Quality*, pages 225–231, Cham. Springer International Publishing.

Hsieh, O. (2016). Human computer interaction and data visualization. *Advanced Writing: Pop Culture Intersections*, pages 1–24.

Huang, D., Tory, M., Adriel Aseniero, B., Bartram, L., Bateman, S., Carpendale, S., Tang, A., and Woodbury, R. (2015). Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):420–433. DOI: 10.1109/TVCG.2014.2359887.

Ismail, N. A. and Zainal Abidin, W. (2016). Data Scientist Skills. *IOSR Journal of Mobile Computing Application*, 03(04):52–61. DOI: 10.9790/0050-03045261.

ISO (1998). Iso 9241-11:1998(en). URL: https://www.iso.org/obp/ui/iso:std:iso:9241:-11:ed-1:v1:en.

ISO/IEC (2011). ISO/IEC 25010:2011, systems and software engineering — systems and software quality requirements and evaluation (square) — system and software quality models.

Janssen, N. (2022). The Data Science Talent Gap: Why It Exists And What Businesses Can Do About It. *Forbes Technology Council*.

Ji, S., Li, Q., Cao, W., Zhang, P., and Muccini, H. (2020). Quality assurance technologies of big data applications: A systematic literature review. *Applied Sciences (Switzerland)*, 10(22):1–31. DOI: 10.3390/app10228052.

Kandari, J., Jones, E. C., Nah, F. F. H., and Bishu, R. R. (2011). Information quality on the World Wide Web: Development of a framework. *International Journal of Information Quality*, 2(4):324–343. DOI: 10.1504/IJIQ.2011.043784.

Knaflic, C. N. (2019). *Storytelling com dados: Um guia sobre visualização de dados para profissionais de negócios*. Alta Books, 2 edition.

Kumar, S. and Jakhar, M. (2010). Understanding user evaluation of Information Quality Dimensions in a digitized world. *Proceedings of Production and Operations Management Society*, 1:1–10.

Law, D. (2020). *Proteção De Dados Do Brasil Lgpd Conhecendo a Lei De Dados Do Brasil*.

Lebo, M. S., Sutti, S., and Green, R. C. (2016). "Big data" gets personal. *Science Translational Medicine*, 8(322):322fs3–322fs3. DOI: 10.1126/scitranslmed.aad9460.

Lee, Y. W., Strong, D. M., Kahn, B. K., and Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management*, 40(2):133–146.

DOI: 10.1016/S0378-7206(02)00043-5.

Lieshout, M. and Kool, L. (2008). Privacy implications of RFID: An assessment of threats and opportunites. In *The Future of Identity in the Information Society*, pages 129–141. Springer US, Boston, MA. DOI: 10.1007/978-0-387-79026-8_9.

Lopes, F. S., da Silva, L. A., and Breternitz, V. J. (2017). Data Scientists: a Study on Skills and Formation. *Proceedings of the 14th CONTECSI International Conference on Information Systems and Technology Management*, 14(May):2075–2081. DOI: 10.5748/9788599693131-14contecsi/rf-4642.

Lopes, L. A., Pinheiro, E. G., Zaina, L. A. M., and Álvaro, A. (2016). A interdisciplinaridade entre a Interação Humano Computador e os Métodos Ágeis na visão dos estudantes. *Anais do VII Workshop sobre Educação em IHC - WEIHC 2016*, pages 7–12.

Mashhadi, A., Kawsar, F., and Acer, U. G. (2014). Human data interaction in iot: The ownership aspect. In *2014 IEEE World Forum on Internet of Things (WF-IoT)*, pages 159–162. DOI: 10.1109/WF-IoT.2014.6803139.

Maus, G. (2015). Decoding, hacking, and optimizing societies: Exploring potential applications of human data analytics in sociological engineering, both internally and as offensive weapons. *Proceedings of the 2015 Science and Information Conference, SAI 2015*, pages 538–547. DOI: 10.1109/SAI.2015.7237195.

McAuley, D., Mortier, R., and Goulding, J. (2011). The Dataware manifesto. *2011 3rd International Conference on Communication Systems and Networks, COMSNETS 2011*. DOI: 10.1109/COMSNETS.2011.5716491.

McGill, M. M., Decker, A., and Abbott, Z. (2018). Improving research and experience reports of pre-College computing activities: A gap analysis. *SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 2018-January(February):964–969. DOI: 10.1145/3159450.3159481.

Ministério da Educação Superior - Conselho Nacional de Educação (2016). Resolução nº 5, de 16 de Novembro de 2016. 2016:1–9.

Morrow, J. (2021). *Be Data Literate: The Data Literacy Skills Everyone Needs to Succeed*. Kogan Page.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., and Crowcroft, J. (2015). Human-Data Interaction: The Human Face of the Data-Driven Society. DOI: 10.2139/ssrn.2508051.

Mortier, R., Haddadi, H., Henderson, T., Mcauley, D., Crowcroft, J., and Crabtree, A. (2016). Human-Data Interaction. *The Encyclopedia of Human-Computer Interaction*, pages 1–48.

Murmann, P. and Fischer-Hübner, S. (2017). Tools for Achieving Usable Ex Post Transparency: A Survey. *IEEE Access*, 5:22965–22991. DOI: 10.1109/ACCESS.2017.2765539.

Neil, T. (2014). *Mobile Design Pattern Gallery*. OReilly, 2 edition.

Peek, S. (2023). Big Data, Big Problem: Coping With a Shortage of Talent in Data Analysis. *business.com*.

Pekala, S. (2017). Privacy and user experience in 21st century library discovery. *Information Technology and Libraries*, 36(2):48–58. DOI: 10.6017/ital.v36i2.9817.

Rocha, H. V. and Baranauskas, M. C. C. (2003). *Design e Avaliação de Interfaces Humano-Computador*. Instituto de Computação - Universidade Estadual de Campinas.

Rogers, Y., Sharp, H., and Preece, J. (2013). *Design de interação: além da interação humano-computador*. Bookman, 3 edition.

Rutgers School of Arts and Science (2022). Data interaction and visual analytics.

Santos, P., Salgado, L., and Viterbo, J. (2018). Assessing the Communicability of Human-Data Interaction Mechanisms in Transparency Enhancing Tools. *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems*, 15:897–906. DOI: 10.15439/2018f174.

Sindiy, O., Litomisky, K., Davidoff, S., and Dekens, F. (2013). Introduction to information visualization (InfoVis) techniques for model-based systems engineering. *Procedia Computer Science*, 16(December 2013):49–58. DOI: 10.1016/j.procs.2013.01.006.

Swinburne University of Technology (2022). Reserach programa in human-data interaction.

Taibi, D., Fernandez-Sanz, L., Pospelova, V., Leon-Urrutia, M., Marjanovic, U., Splendore, S., and Urbsiene, L. (2021). Developing Data Literacy Competences at University: The experience of the DEDALUS project. *2021 1st Conference on Online Teaching for Mobile Education, OT4ME 2021*, pages 112–113. DOI: 10.1109/OT4ME53559.2021.9638912.

Toledo, M. D. E. (2020). Lei Geral de Proteção de Dados. um guia completo.

Tom, J., Sing, E., and Matulevičius, R. (2018). Conceptual representation of the GDPR: Model and application directions. *Lecture Notes in Business Information Processing*, 330(January):18–28. DOI: 10.1007/978-3-319-99951-7_2.

United Nations Development Group (2017). UNDG Guidance Note on Big Data for Achievement of the 2030 Agenda : Data Privacy, Ethics, and Protection. page 16.

Vivacqua, A. S., França, J. B. d. S., and Dias, A. F. d. S. (2019). Promoting Data Literacy in Brazilian Schools. *Proceedings of CLIHC '19*, 4:1–4. DOI: 10.1145/3358961.3359003.

Wroblewski, L. (2011). *Mobile First*. DOI: 978-1-937557-02-7.

Zorzo, A. F., Nunes, D., Ecivaldo, S., and Martins, S. (2017). *Referenciais de Formação para os Cursos de Graduação em Computação 2017*. Sociedade Brasileira de Computação.