Frameworks and Methods for Evaluating Open Data Platforms Under the Lens of Human-Data Interaction: A Systematic Review

Luana B. Mendes[1], Cristiano Maciel[1 2], Rita Berardi[3], Eunice P. dos Santos Nunes[1 2]

[1] Graduate Program in Applied Computing (PPGCOMP) – Federal University of Mato Grosso (UFMT) – Cuiabá – MT – Brazil

[2] Institute of Computing (IC) – Federal University of Mato Grosso (UFMT) – Cuiabá – MT – Brazil

[3]Federal University of Technology – Paraná (UTFPR)

luana-bm@hotmail.com,{cristiano.maciel, eunice}@ufmt.br,
ritaberardi@utfpr.edu.br

Abstract. Introduction*: This systematic review analyzes the literature on Human-Data Interaction (HDI) within the context of open data platforms.* ***Objective****: To investigate HDI definitions and challenges, evaluate approaches for assessing platform quality and complex data visualizations, and understand user needs, including those of non-experts.* ***Methodology****: A predefined protocol guided a search in digital databases, followed by a rigorous analysis of 25 selected studies.* ***Results****: The review reveals a lack of frameworks for evaluating open data platforms. It also highlights significant interaction challenges, such as difficulties in data interpretation and the need for better support for non-specialized users.*

Keywords *Human-Data Interaction; Open Data Platforms; Evaluation.*

1. Introduction

The exponential growth in the amount of data generated by society has transformed the ability to handle large volumes of information, originating from diverse sources and in different formats, into a highly valued skill [Isotani and Bittencourt 2015]. In parallel, the phenomenon known as the "data deluge" emerges [Coleti et al. 2024], characterized by the abundance of data that is often not publicly available or is not presented in a structured manner, hindering its interpretation and use [Isotani and Bittencourt 2015; Vasconcelos et al. 2020], highlighting the scarcity of open data [Janssen et al. 2012]. Faced with this scenario, the efforts of companies, governments, and research institutions to make open data available and develop web-based technologies have intensified, aiming to foster an ecosystem of data production and consumption that promotes knowledge discovery and adds value to information accessible through the internet [Natvig et al. 2021].

In this context, the concept of Open Data emerges, defined by the Open Definition as data that can be freely used, reused, and redistributed, but with the requirement of citing the original source and sharing under the same licenses, being published in an open and readable format, both for people and for machines. The opening of data brings benefits such as the possibility of reusing existing data, tools, and libraries, combining data with other sources, and increasing the value of information. However, for this to happen, it is necessary to structure the data, and this is possible using methods such as Ontologies

[Isotani and Bittencourt 2015]. There are several methodologies for building ontologies, such as DILIGENT (which considers aspects such as decentralization, partial autonomy, continuous iteration, and the participation of non-experts [Pinto et al. 2009]), and HCOME (Human-Centered Ontology Engineering Methodology [Kotis and Viyris 2006]), which are inspired by Human-Computer Interaction (HCI) and have greater user involvement, considering collaboration and negotiation activities in the creation process [Hornung et al. 2015].

The difficulty in understanding and interacting with the growing volume of available data limits the ability of individuals and organizations to fully explore the informative potential of these resources, compromising the effectiveness in formulating strategies and making informed decisions [Victorelli et al., 2019]. This interaction between users and data has led to the emergence of the research area known as Human-Data Interaction (HDI), which generally deals with the interaction between human beings, datasets, and analysis [Mortier et al. 2014]. While HCI includes the processes of rigorously designing, implementing, and evaluating computer systems, with a central focus on delivering systems that are inherently easy to use and provide greater user satisfaction, HDI focuses on understanding how human beings interact with data in information systems, being an interdisciplinary area of knowledge [Coleti et al. 2024]. Several studies in the field of HCI seek to explain the objectives and boundaries of the field, introducing the problem, defining concepts, and presenting challenges related to the topic [Mortier et al. 2014; Hornung et al. 2015], clarifying that designing software applications that facilitate the interaction of users with different profiles with data is not a trivial task [Victorelli and Reis 2020].

According to Coleti et al. (2024), it is crucial to consider the ability to manipulate, analyze, and make decisions based on complex data; the need to provide personalized and understandable data and information; and the development of resources that assist individuals in accessing, understanding, and using this data. Furthermore, it is important to create an interaction flow that favors the development of analytical reasoning and enables decision-making based on this understanding, facilitating the comprehension of data by users [Victorelli and Reis 2020].

It is worth noting that the "non-expert users" mentioned in some methodologies refer to end-users who do not possess advanced technical knowledge in computer science, interface design, or areas related to this investigation, similar to Tylosky et al. (2025). Additionally, the Visualization of complex data was considered, characterized by typically being extensive, in diverse formats, often unstructured, and continuously generated at high speed, demanding new tools and techniques for processing, analysis, and interpretation [Kitchin and McArdle 2016], in order to give greater breadth to the research.

Given this overview, this article presents the results of a Systematic Literature Review (SLR) concerning Human-Data Interaction. This review is dedicated to identifying frameworks, metrics, methods, methodologies, and heuristics applicable to the evaluation of open data platforms from the perspective of HDI. Additionally, the literature was investigated for interaction challenges and user needs, including non-expert users, as well as the improvements conceived within HDI to mitigate these challenges and difficulties.

These investigative efforts are part of a project aimed at developing an open data platform called ELLAS (Equality in Leadership for Latin-American STEM), whose scope includes evaluating the quality of the platform itself. It is worth noting that this study

aligns with the Grand Challenges of HCI [Pereira et al. 2024], more specifically the GranDIHC-BR 2025-2035 - GC5 - Human-Data Interaction, Data Literacy, and Usable Privacy [Coleti et al. 2024], which addresses HDI as a research field still in its initial stages (with publications since 2010) and points to the existence of a series of challenges related to human factors, which should be the focus of research and development, and align with the proposition of this study.

In this context, this article presents the initial results of this evaluation process, initiated through an SLR, with the objective of identifying quality criteria in open data platforms. These criteria will serve as the basis for the development of an artifact that will assist in the analysis and evaluation of the quality of this type of platform.

2. Methodology

SLR is a research method characterized by a structured process, the objective of which is to identify, select, evaluate, and synthesize the relevant scientific literature to answer specific research questions that must be previously defined [Kitchenham 2007; Grant and Booth 2009]. The objective of the SLR is to contextualize Human-Data Interaction through publications from the last five years in the context of open data platforms. The main focus lies in understanding the primary challenges and needs of users, seeking to identify methods, methodologies, metrics, heuristics, and/or frameworks for the evaluation of open data platforms found in the literature.

Initially, a protocol was developed to guide this systematic review, serving as a guideline for the execution of all stages of the process. The SLR sought to answer the following research questions:

1. How does the scientific literature define Human-Data Interaction, particularly within the domain of open data platforms?
2. What frameworks, models, methods, metrics, or heuristics exist in the literature to evaluate the quality of open data platforms or complex data visualization systems?
3. What are the main interaction challenges and specific needs of users (including non-experts) when using open data platforms, as reported in the literature?

The selection of studies was conducted by an IHC researcher based on the following inclusion criteria: publications from the last five years (2020 - 2025), studies published in Portuguese or English, and studies that answer the research questions. Duplicate studies, short articles (with four pages or less), grey literature, and studies outside the scope of HDI and open data platforms were excluded. The literature search was conducted in the following databases: ACM Digital Library (http://portal.acm.org), IEEE Digital Library (http://ieeexplore.ieee.org), Scopus (http://www.scopus.com), and Springer Link (http://link.springer.com), given that they represent the main databases of publications used in various international conferences and in multiple countries. The conduct of the SLR was supported by the Parsifal tool (available at: https://parsif.al), which assisted in defining the protocol, organizing the stages of the review process, and managing references, in order to contribute to greater rigor, transparency, and reproducibility of the study.

The following search string was applied in the databases, adapted according to the specificities of each database: ("*Human-Data Interaction*" **OR** *HDI* **OR** "*Data Interaction*") **AND** ("*Open Data Platform**" **OR** "*Open Data Portal**") **AND** (*Evaluat** **OR** *Assess** **OR** *Measur** **OR** *Framework** **OR** *Model** **OR** *Method** **OR** *Metric** **OR**

*Heuristic**) **AND** (*Quality* **OR** *Usability* **OR** *Accessibility* **OR** "*User Experienc*e" **OR** *UX*), resulting in the identification of 55 records.

The relevant data from the studies were extracted using a standardized form, with information on the authors, publication date, number of pages, publication venue, number of citations, and abstract. After applying the exclusion criteria (removal of duplicates, grey literature, and publications in languages other than English and Portuguese), 38 records were selected for the next stage, as shown in Table 1.

Subsequently, an analysis of the abstracts of the pre-selected articles was conducted. The inclusion criterion at this stage was the relevance of the abstract's content to the scope of this SLR, evaluated based on the following thematic axes: HDI, evaluation of open data platforms, complex data visualization, design of open data platforms, and challenges inherent to the use of such platforms or the visualization of complex data.

Table 1 - Research protocol results

Database	Results	Retained Articles
Springer	12	6
IEEE	12	8
ACM	10	10
Scopus	21	14

For this purpose, the quality assessment questions available in the Parsifal tool were applied, which are defined criteria to evaluate the reliability, relevance, and methodological rigor of the articles included in the SLR. These questions help to identify whether the selected studies are sufficiently robust to support the review's conclusions. These questions were defined during the planning of the review.

QQ1: Does the study propose an evaluation of HDI for open data platforms?

QQ2: Does the study have a detailed structure for conducting an HDI evaluation?

QQ3: Does the study identify HDI challenges?

QQ4: Does the study identify HDI problems in open data platforms?

QQ5: Does the study outline methods for improving human interaction with data?

Based on the abstracts of each article, the quality questions were answered with "Yes," "No," or "Partially." This allowed for the calculation of a quality score for each study, assisting in the screening and analysis of the data. After this classification, all articles that received more than two "No" answers or did not have at least two "Yes" answers were excluded.

This quality process was conducted by an HCI researcher and reviewed by two others. At the end of the process, 10 articles remained, as detailed in Table 2, showing the number of articles per database that met this criterion and were considered in full for analysis, thus forming the final corpus of this review. It should be noted that the citation frequency of the articles in other publications was not a criterion for inclusion or

exclusion, since this review aims to analyze exclusively recent publications, covering the period from 2020 to March 2025.

In addition to the results obtained from the electronic searches conducted in the four cited libraries, a manual search was implemented. That is, during the analysis of the abstracts, the references of the pre-selected articles were verified. This process, also known as "snowballing" [Runeson 2009], allowed for the identification of other studies of potential interest, whose abstracts were evaluated for possible inclusion in this review.

Table 2 - Articles for full-text reading

Paper	Database	General Overview
[Tylosky et al. 2025]	Scopus	Focuses on design practices for data exploration through visualization for non-expert audiences, mentioning open data portals as a context.
[Liang et al. 2025]	Scopus	Does not propose a new HDI evaluation but applies existing models (D&M, DEWEM) to evaluate the continued use and value creation in OGD, which are central to HDI.
[Koesten et al. 2020]	Scopus	Implicitly identifies problems in how datasets are presented/summarized.
[Barcellos et al. 2024]	ACM	Discusses problems such as difficulty in interpretation, the need for better data presentation, and the lack of traceability as barriers to the effective use of data by citizens.
[Vasconcelos et al. 2020]	Scopus	Highlights data communicability issues.
[Degbelo 2020]	ACM	Outlines user needs that serve as targets for improvement but does not detail the methods to address these needs.
[Tylosky et al. 2023]	IEEE	Identifies general design challenges for exploration systems.
[Xiao et al. 2023]	Scopus	Proposes an interaction model for HDI in OGD.
[Puussaar et al. 2022]	ACM	Identifies problems in how the smart city sensor infrastructure excludes citizens or isolates their data, not specifically in ODPs.
[Barcellos et al. 2022]	ACM	Develops and proposes a model with 8 dimensions to formally define the interpretability of data in open portals.

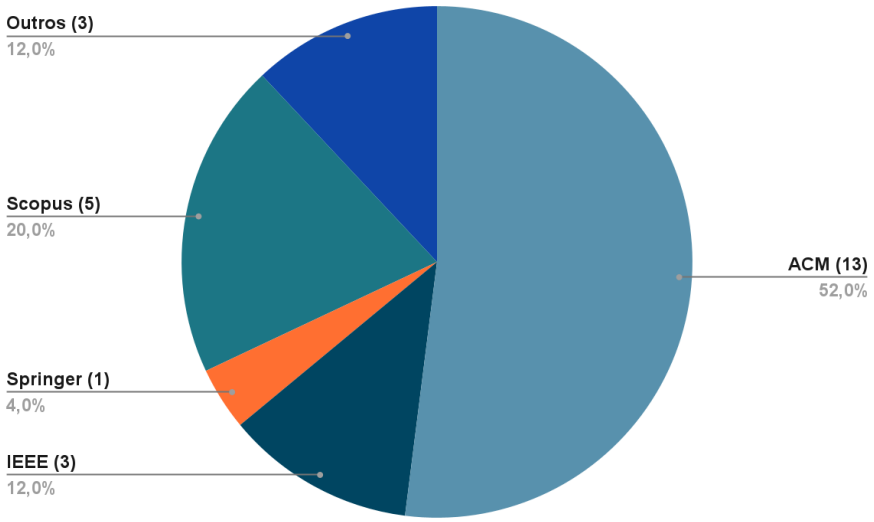
The adoption of this complementary strategy is based on the recognition that electronic searches, although essential and widely recommended by the methodological literature, may be insufficient. Furthermore, manual searching of references is advocated by researchers as a means of broadening the scope of the retrieval of pertinent studies [Kitchenham 2007]. The articles included in the review through this manual approach are listed in Table 3, and since it was a manual search, the inclusion criterion that considered only studies published in the last five years was not applied.

Table 3 - Articles selected through manual search

Paper	Publication year	Database
[Cantador et al. 2021]	2021	ACM
[Barcellos et al. 2018]	2018	IEEE
[Sampaio et al 2022]	2022	Springer
[Xiao et al. 2022]	2022	ACM
[Máchová e Lněnička 2017]	2017	SciELO
[Coleti et al. 2024]	2024	ACM
[Oliveira et al. 2023]	2023	SOL
[Barcellos et al. 2023]	2023	ACM
[Victorelli e Reis 2020]	2020	ACM
[Flausino et al. 2024]	2024	SOL
[Hsu et al. 2021]	2021	ACM
[Lee et al. 2020]	2020	IEEE
[Xiao et al. 2020]	2020	ACM
[Tylosky et al. 2023]	2023	ACM
[Ferreira et al. 2024]	2024	ACM

Thus, the final sample of studies included in this SLR totals 25 publications, as illustrated in Figure 1.

Figure 1 - Articles under analysis in this review



The 15 manually selected studies were integrated into the analysis table, along with the 10 articles selected in the automated search. In this table, the studies were also classified according to their quality, based on the five quality questions. The final result of the quality analysis of all evaluated articles (automatic and manual) is detailed in Table 4 where “N” stands for “No”, “S” for “Yes” and “P” for “Partially.

Table 4 - Quality classification of the selected studies

Paper	Quality Questions				
	QQ1	QQ2	QQ3	QQ4	QQ5
[Tylosky et al. 2025]	N	N	S	S	S
[Liang et al. 2025]	P	P	S	S	S
[Koesten et al. 2020]	N	N	S	P	S
[Barcellos et al. 2024]	N	N	S	S	S
[Vasconcelos et al. 2020]	P	S	S	S	P
[Degbelo 2020]	N	N	S	S	P
[Tylosky et al. 2023]	N	N	S	P	S
[Xiao et al. 2023]	N	N	S	S	S
[Puussaar et al. 2022]	N	N	S	P	S
[Barcellos et al. 2022]	P	N	S	S	S
[Cantador et al. 2021]	N	N	S	S	S
[Barcellos et al. 2018]	N	P	S	S	S
[Sampaio et al 2022]	N	N	S	S	P
[Xiao et al. 2022]	N	N	S	S	P
[Máchová e Lněnička 2017]	P	P	S	S	S
[Coleti et al. 2024]	N	N	S	P	S
[Oliveira et al. 2023]	N	N	S	P	S
[Barcellos et al. 2023]	N	N	S	S	S
[Victorelli e Reis 2020]	N	N	S	P	S
[Flausino et al. 2024]	N	N	S	S	P
[Hsu et al. 2021]	N	N	S	S	S
[Lee et al. 2020]	N	N	S	P	S
[Xiao et al. 2020]	N	N	S	S	P
[Tylosky et al. 2023]	N	N	S	P	S
[Ferreira et al. 2024]	N	N	S	S	S

After consolidating the selected articles and the quality assessment (QQ), each study was classified according to the relevance of its content to the three proposed RQ (research questions), using the following parameters: "High," "Moderate," "Low," and "None." The result of this classification is detailed in Table 5.

Table 5 - Relevance classification of the selected studies

Paper	Research Questions		
	RQ1	**RQ2**	**RQ3**
[Vasconcelos et al. 2020]	High	High	High
[Barcellos et al. 2022]	High	High	High
[Victorelli e Reis 2020]	High	High	High
[Tylosky et al. 2025]	Moderate	Low	High
[Liang et al. 2025]	Low	High	Moderate
[Koesten et al. 2020]	High	Moderate	High
[Barcellos et al. 2024]	High	Moderate	High
[Tylosky et al. 2023]	High	Moderate	High
[Xiao et al. 2023]	High	Moderate	High
[Barcellos et al. 2018]	Moderate	High	High
[Flausino et al. 2024]	Moderate	High	High
[Coleti et al. 2024]	High	Moderate	High
[Oliveira et al. 2023]	Moderate	High	High
[Cantador et al. 2021]	Low	High	High
[Barcellos et al. 2023]	Moderate	Moderate	High
[Degbelo 2020]	Moderate	Moderate	High
[Tylosky et al. 2023]	Moderate	Moderate	High
[Xiao et al. 2020]	Moderate	Moderate	High
[Máchová e Lněnička 2017]	Low	High	Moderate
[Lee et al. 2020]	Moderate	Low	High
[Xiao et al. 2022]	Moderate	None	High
[Sampaio et al 2022]	Low	High	Low
[Hsu et al. 2021]	Moderate	Moderate	Moderate
[Puussaar et al. 2022]	Moderate	Low	Moderate
[Ferreira et al. 2024]	None	Low	High

3. Results and Discussion

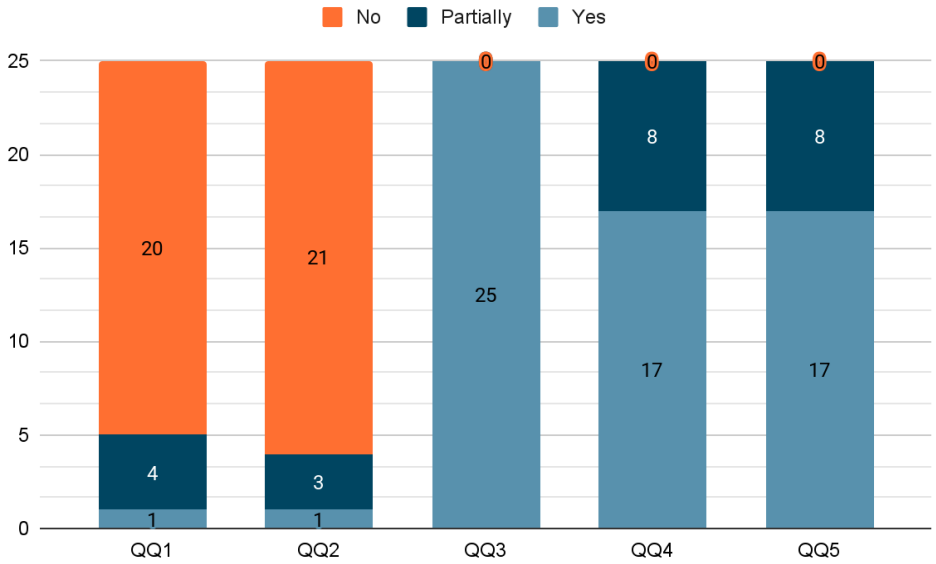
This section is dedicated to presenting the results of the SLR through a synthesis of the 25 studies selected for this investigation.

3.1 Quantitative Evaluation of the Studies

During the quality evaluation process, it was found that the majority of the studies do not present a specific methodology, method, metrics, framework, or heuristic for the evaluation of open data platforms, highlighting a gap in the publications on the subject. The quality analysis of the included studies, detailed in Table 4, is presented from a

quantitative perspective in Figure 2. It is noted in quality question QQ1 that only one of the studies included in the review does not propose an HDI evaluation for open data platforms ("Yes" answer for QQ1), while other authors who bring proposals for quality evaluation do not focus on HDI or open data platforms and were therefore classified as "Partially."

Figure 2 - Articles under analysis in this review



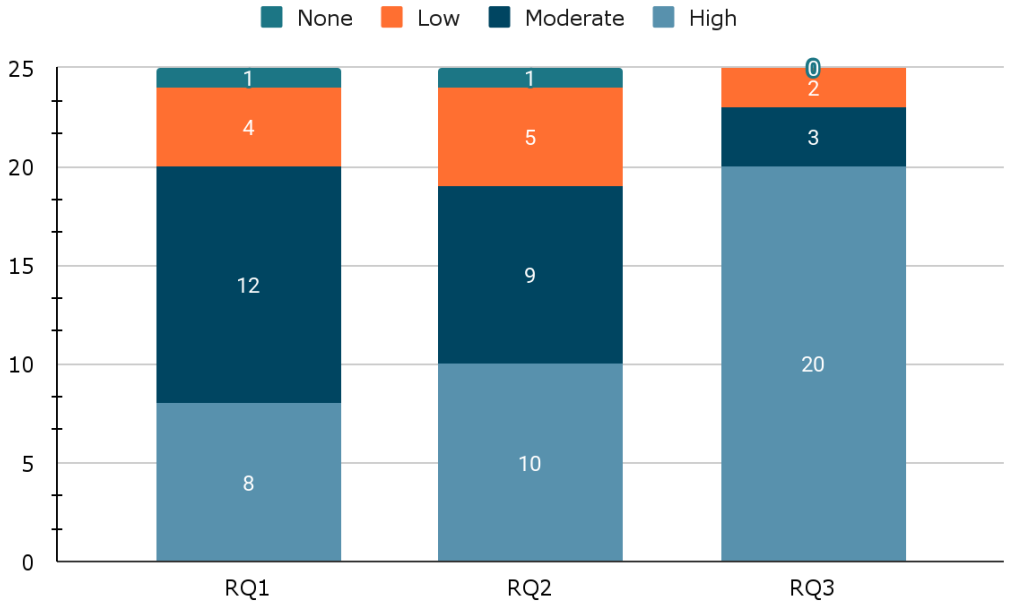
The analysis revealed that eight (five from the automated search and three from the manual search) of the 25 examined studies presented high relevance for understanding and potentially elucidating the first research question: "How does the scientific literature define HDI, particularly within the domain of open data platforms?". In contrast, five studies showed low or no relevance to this question (one from the automated search and four from the manual search).

Regarding the second research question, "What frameworks, models, methods, metrics, or heuristics are presented in the literature for evaluating the quality of open data platforms or complex data visualization systems?", 10 results of high relevance were identified (three from the automated search and seven from the manual search). In contrast, six studies were classified as having low or no relevance to this question.

In conclusion, 20 studies (eight from the automated search and 12 from the manual search) were considered of high relevance to the third research question: "What are the main interaction challenges and specific needs of users (including non-experts) when using open data platforms, as reported in the literature?".

In contrast, only two studies were classified as having low relevance, as illustrated in Figure 3.

Figure 3 - Articles under analysis in the review



Furthermore, it was found that three studies were classified as having "high" relevance for all RQs, as detailed in Table 6.

Table 6 - Articles of high relevance for the study

Paper	Search	Year	Published on
[Barcellos et al. 2022]	Automatic	2022	Information Systems
[Victorelli e Reis 2020]	Manual	2020	19th Brazilian Symposium on Human Factors in Computing Systems
[Vasconcelos et al. 2020]	Automatic	2020	26th Americas Conference on Information Systems

Each subsequent subsection presents the title corresponding to the research question and a brief analysis of the selected articles, highlighting the main findings for the addressed question.

3.2 How does the scientific literature define Human-Data Interaction (HDI), particularly within the domain of open data platforms?

As mentioned in subsection 3.1, eight studies were considered highly relevant to the theme of the first research question, allowing us to observe that the definition of HDI presented, even over the years, shows consistency and convergence.

It was found that, among the studies considered highly relevant to the first research question, three do not mention the terms "open data," as shown in Table 7, but allude to "data visualization." This quantitative analysis reveals that, predominantly, "data visualization," "open data," and "human-data interaction" constitute interrelated concepts

in the selected studies, making it pertinent to also consider "data visualization" when investigating "Open Data," since open data platforms also represent a form of complex data visualization.

Table 7 - Term mentions per study

Paper	"IHD", "*HDI*", "interação humano-dados" or "*Human-data interaction*"	"*open data*" or "dados abertos"	"visualização de dados" or "*data visualization*"
[Xiao et al. 2023]	47	24	3
[Vasconcelos et al. 2020]	41	68	17
[Koesten et al. 2020]	12	6	2
[Barcellos et al. 2024]	6	70	10
[Barcellos et al. 2022]	5	85	0
[Coleti et al. 2024]	80	0	7
[Victorelli e Reis 2020]	61	0	6
[Tylosky et al. 2023]	7	0	7

According to the analyzed works, HDI is conceptualized as an interdisciplinary research field dedicated to understanding the interaction between human beings and data in information systems [Koesten 2020]. This interaction encompasses the ability to manipulate and analyze complex data, grounding decision-making; providing personalized and intelligible data; assisting individuals in accessing and understanding data and its behavioral impacts; managing diverse databases; and controlling events related to this data [Coleti et al. 2024]. Furthermore, the studies emphasize the interdisciplinary nature of HDI, integrating domains such as computing (notably HCI and Complex Data Visualization - raw and structured), statistics, sociology, psychology, and behavioral economics [Vasconcelos et al. 2020; Xiao et al. 2023].

Barcellos et al. (2022), based on the concepts presented by Mortier et al. (2014), conceptualize HDI as the analysis of decisions and actions, both individual and collective, carried out while users of online systems or as subjects of data collection, with an emphasis on personal data and open data. The authors discuss the three main aspects for data interaction, originally listed by Mortier et al. (2014):

1. Legibility: Concerning the process of making data and analysis algorithms transparent and understandable to the public, both in relation to the data itself and its processing.
2. Agency: Related to the power to manipulate data and the systems that process it, aiming to give citizens the ability to control, inform, and infer accurate data.

3. Negotiability: Referring to the dynamic relationships that emerge around data and how individual understanding and attitudes evolve over time. It involves supporting individuals in re-evaluating their decisions as contexts change. Additionally, it relates to the formation of social norms and applicable legal and regulatory structures for data.

In the research by Barcellos et al. (2024), HDI is not conceptualized as explicitly, but the authors emphasize the relevance of HDI for open data portals, with a specific focus on open government data portals. The authors address various initiatives present in the literature that aim to fill the gap identified in the study: the absence of efficient architectures, frameworks, processes, and methodologies to identify, map, develop, and plan resources that optimize HDI. Works are mentioned that seek to promote more efficient interfaces, offering distinct functionalities that facilitate the understanding of data by different user profiles, as well as data visualization options.

Due to their focus on the design of systems for data visualization, Victorelli and Reis (2020) base their work not only on specific HDI authors but also on the literature from other related areas, such as HCI and Information Visualization. Drawing on the work of Elmqvist (2011) and other proprietary research, Victorelli et al. (2019), the authors define HDI as the term used to address the human manipulation, analysis, and interpretation of extensive, unstructured, and complex datasets, encompassing different aspects such as personal data, unstructured datasets, transparency, trust, privacy, and embedded interaction environments. The authors also reinforce that HDI comprises the intersection of areas such as HCI and Information Visualization, as do other studies analyzed in this review, and that these areas should be approached with a focus to more accurately achieve HDI-related objectives, emphasizing the need to reuse, combine, and harmonize existing practices in related domains. In this sense, it becomes imperative to establish an improved understanding of how to design to favor HDI.

Finally, Tylosky et al. (2023) do not address HDI as an autonomous research field but rather as a topic to be investigated within the scope of HCI and data visualization, which is why the study does not present a formal definition of HDI. However, the work focuses on the design of data visualizations aimed at non-expert users, in line with the concept defined in the introduction of this SLR. The authors also point out that the lack of consideration for non-expert audiences during the design of data exploration systems can constitute an obstacle for novice users of these systems and discourage the engagement of non-expert users with open data. This observation is based on recent literature reviews on HDI systems and user interfaces in the field of HCI, in which the focus has been on usability in general, without specific consideration of the target users of these systems and interfaces.

3.3 What frameworks, models, methods, metrics, or heuristics exist in the literature to evaluate the quality of open data platforms or complex data visualization systems?

Following the quality analysis of the included studies—the step subsequent to the inclusion and exclusion process—a gap was identified in the literature concerning the existence of frameworks, models, methods, metrics, or heuristics directed at the evaluation of open data platforms. As presented in Table 4 and Figure 2, only one of the studies selected for analysis in this review was qualified with "Yes" for quality question QQ1 ("Does the study propose an evaluation approach for HDI on open data platforms?"), and only five were classified as "Partially" (a category attributed to works that presented

an evaluation structure but not specifically for HDI). The remaining 19 studies analyzed in this stage were categorized as "No" for QQ1, indicating the absence of an HDI evaluation proposal, or any other type of evaluation, for open data platforms.

Furthermore, two other studies presented evaluation proposals—being classified as "Yes" or "Partially" for QQ2 ("Does the study present a detailed method for HDI evaluation?")—specifically directed at data visualization but not in the context of open data.

Thus, a total of six studies address, at least partially, the existence of frameworks, models, methods, metrics, or heuristics for the evaluation of data platforms (not necessarily HDI—classified as "partially") or data visualization, out of the 25 studies analyzed, as presented in Table 5.

Moreover, two other studies were considered relevant to RQ2. Although they do not present a formal framework for evaluating the quality of open data platforms or data visualization (they were classified as "No" for both QQ1 and QQ2), these works offer specific contributions on how to evaluate particular aspects of these platforms or specific open data platforms.

All articles of interest for answering RQ2 are listed in Table 8 and will be analyzed in this section.

Table 8 - Relevant studies on data and/or data platform evaluation

Paper	QQ1	QQ2	RQ2
[Vasconcelos et al. 2020]	Yes	Yes	High relevance
[Liang et al. 2025]	Partially	Partially	High relevance
[Máchová e Lněnička 2017]	Partially	Partially	High relevance
[Barcellos et al. 2022]	Partially	No	High relevance
[Barcellos et al. 2018]	Partially	Partially	High relevance
[Sampaio et al 2022]	Partially	No	High relevance
[Oliveira et al. 2023]	No	No	High relevance
[Flausino et al. 2024]	No	No	High relevance

It is worth noting that, in the eight mentioned articles, evaluations carried out through frameworks, checklists, and heuristics were identified; however, no metrics, methods, and/or evaluation methodologies were found.

3.3.1 Studies containing evaluation methods, methodology, framework, metrics, or heuristics

The research by Liang et al. (2025), Máchová and Lněnička (2017), Barcellos et al. (2018), and Barcellos et al. (2022) present an evaluation framework, but without a focus on HDI.

Vasconcelos et al. (2020) investigate the communicability difficulties faced by non-expert users throughout the entire open data manipulation flow (collection from the

portal, analysis/visualization in an external tool). The authors employ the Semiotic Inspection Method (SIM), a qualitative inspection method to evaluate the communicability of interactive systems, in the analysis of the interfaces of a Brazilian data portal (ISP Dados) and the Power BI tool. This inspection is conducted in five stages detailed by the authors, in which it is emphasized that, although the study focuses on non-expert users, a method was chosen in which experts act as user advocates when investigating potential communicability problems in the interface and their consequences for HCI.

The results of the SIM are triangulated with data from an experiment in which students performed tasks in the flow and answered questionnaires, whose responses were analyzed qualitatively to identify the main communication challenges, such as non-intuitive interfaces and inconsistent data. Vasconcelos et al. (2020) also mention difficulties encountered during the study, such as non-intuitive interfaces, lack of usage tutorials (user manuals), excessive difficulties in accessing platforms, and inconsistent data, showing concern for HDI.

Barcellos et al. (2018) analyze the quality of data visualizations in OGD portals, highlighting that many are ambiguous or confusing, hindering understanding, especially for non-expert users. Although the study follows HCI principles and does not mention HDI, it proposes an evaluation instrument based on 11 consolidated heuristics from the HCI and Information Visualization literature (covering aspects such as color, layout, information overload, 3D, etc.), aimed at designers with different levels of expertise, to improve the clarity, accuracy, and efficiency of visualizations.

A statistical analysis (using Cohen's Kappa and Pairwise Agreement) indicated that the instrument is suitable and usable by non-experts to identify heuristic violations, although the accuracy in the exact counting of violations may vary, demonstrating the feasibility of heuristic evaluation in this context.

Liang et al. (2025), in turn, evaluate the quality of open data portals (specifically open government data). However, the term "quality" used to describe the evaluation does not apply to HDI but rather to the quality of the data contained in these platforms. The study investigates the determinants of the continued use of open government data portals and their impact on value creation (net benefits and trust in government). To this end, it proposes a theoretical model that integrates self-regulation frameworks, the D&M model (synthesis of the McLean and DeLone models), and the DEWEM (Democratic E-Government Website Evaluation Model). The evaluation process covered two main dimensions: website evaluation (which includes information adequacy, transparency, service quality, security, and citizen engagement) and data quality (which includes data accessibility, completeness, and timeliness), aiming to understand how "website functionality" and "data quality" influence satisfaction and perceived value, which, in turn, impact continued use. The methodology employed involved the analysis of a survey applied to 267 qualified users of OGD (Open Government Data) in China, with the data analyzed using the PLS-SEM (Partial Least Squares Structural Equation Modeling) statistical technique, offering a replicable model.

Analogously to Liang et al. (2025), Máchová and Lněnička (2017) evaluate the quality of governmental open data portals, although without a focus on HDI. The authors propose and validate a national-level benchmarking framework, highlighting the importance of discovering relevant data and creating quality metadata to facilitate access and use by different audiences, including non-experts.

The framework proposed by Máchová and Lněnička (2017) derives from the literature and considers both the functionalities of the portal and the quality of the datasets and metadata. The evaluation is multidimensional, combining content analysis via API with the application of the framework through a questionnaire composed of 28 criteria (all detailed in the study for replication), with possible answers on a five-point Likert scale (1 = "Strongly Disagree," 2 = "Disagree," 3 = "Neutral," 4 = "Agree," 5 = "Strongly Agree"). The questionnaire was initially applied to two professors with extensive experience in quantitative research and, subsequently, to postgraduate students, covering technical, access, participation, and data quality dimensions. The study analyzed 67 national open data portals (from various countries), and the comparison was based on the 2015 GODI (Global Open Data Index) and ODB (Open Data Barometer) rankings, which assess the state of open data in selected countries globally. The methodology employed included accessing the selected portals via API for the analysis of datasets and metadata regarding the date of creation and modification. After the detailed application of the framework, the results indicate that the open data portal with the best quality was that of the United Kingdom, with 129.6 points, while that of Israel obtained the lowest score, with 68.7 points.

The same framework created and validated by Máchová and Lněnička (2017) is replicated by Sampaio et al. (2022) with a focus on the analysis of seven open data platforms on COVID-19 (such as Johns Hopkins, WHO, Brasil IO), demonstrating the applicability of the framework in this new context, also without a focus on HDI.

Barcellos et al. (2022), in turn, focus on the concept of data interpretability in the context of open data portals, seeking to establish a formal definition, without, however, associating it with the quality of the portals or the definitions of HDI. The authors developed an eight-dimensional conceptual model to evaluate the interpretability of open data portals through an analysis of the literature and the NFR (Non-Functional Requirements) framework. A qualitative analysis was also carried out with experts in data analysis, electronic government, and transparency, to understand the definition of interpretability from the perspective of these professionals. Based on these two studies, the authors defined "interpretability" through an eight-dimensional model, each containing different characteristics that must be ensured in the interpretability process for the data to be correctly interpreted.

Barcellos et al. (2022) also investigated technologies and methods associated with the characteristics of interpretability, highlighting the need for computational tools that assist users in understanding governmental open data. The resulting checklist was applied to three portals—European Open Data Portal (data.europa.eu), American Open Data Portal (data.gov), and Brazilian Open Data Portal (dados.gov.br)—providing a replicable methodological basis. The authors conclude that several challenges still persist in these environments.

3.4 What are the main interaction challenges and specific needs of users (including non-experts) when using open data platforms, as reported in the literature?

Of the 25 studies analyzed, 17 were classified with "Yes" for quality question QQ4 ("Does the study identify HDI problems in open data platforms?") and 17 also received "Yes" for QQ5 ("Does the study present clear methods for improving the interaction between humans and data?"). Among these, 11 studies simultaneously met both criteria. However, only 8 studies were considered highly relevant to answer RQ3 — "What are the main

interaction challenges and specific needs of users (including laypersons or non-experts) when using open data platforms, as reported in the literature?", as can be seen in Table 9.

Table 9 - Relevant studies for the identification of challenges in HDI

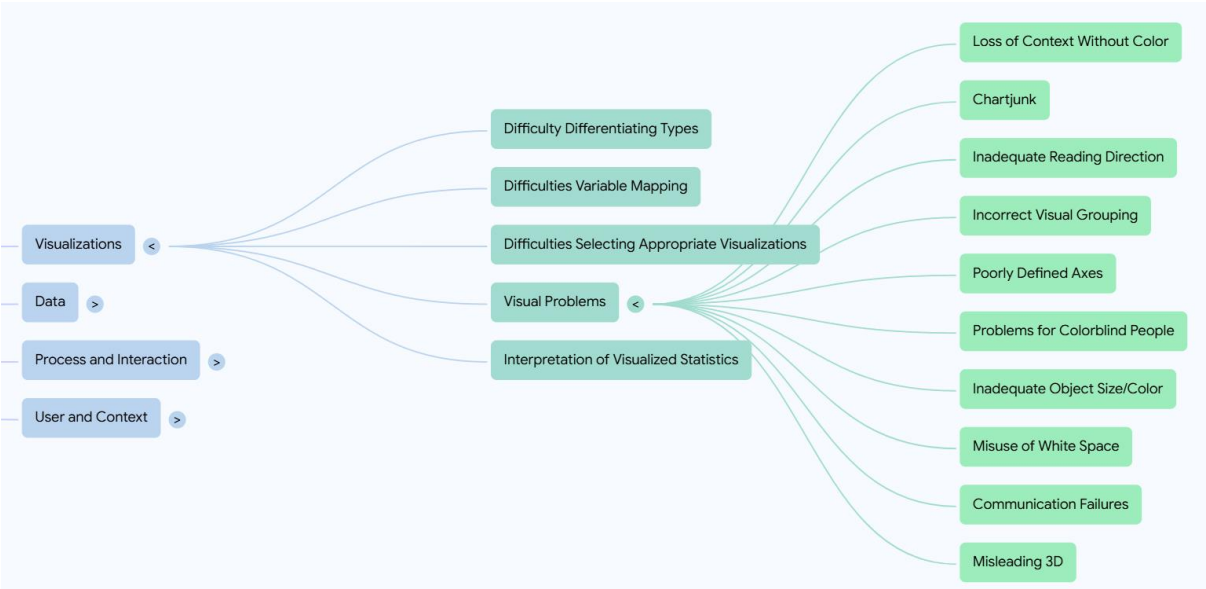
Paper	QQ4	QQ5	RQ3
[Tylosky et al. 2025]	Yes	Yes	High
[Barcellos et al. 2024]	Yes	Yes	High
[Xiao et al. 2023]	Yes	Yes	High
[Barcellos et al. 2022]	Yes	Yes	High
[Cantador et al. 2021]	Yes	Yes	High
[Barcellos et al. 2018]	Yes	Yes	High
[Barcellos et al. 2023]	Yes	Yes	High
[Ferreira et al. 2024]	Yes	Yes	High

Some studies, despite addressing difficulties and/or challenges related to HDI, presented reduced relevance for this study because they dealt with specific platforms or data types, did not directly focus on the context of open data, or did not highlight user interaction with the data, as is the case with Máchová and Lněnička (2017), who present a quality assessment of the API (Application Programming Interface) structure of government open data portals, without focusing on user interaction with the data.

Among the eight main articles pertinent to RQ3, problems and challenges were identified and grouped into four categories: Visualization (Figure 4), Data (Figure 5), Process and Interaction (Figure 6), and User and Context (Figure 7).

Although some elements fit into more than one category, and different authors use distinct terminologies to refer to the same problems, it is observed that many of these challenges are already reported in the literature. Still, despite the fact that the majority of the articles in this SLR address issues related to user interaction with large volumes of data, only the eight most relevant studies—according to the applied quality criteria—were selected for highlighting. Problems related to data visualization, as illustrated in Figure 4, are mentioned by authors such as Ferreira et al. (2024), Tylosky et al. (2025), Barcellos et al. (2018, 2022), which shows that these are recurring and well-documented challenges that still persist in practice.

Figure 4 - Challenges and problems identified in the literature regarding visualizations



Although some elements fit into more than one category, and different authors use distinct terminologies to refer to the same problems, it is observed that many of these challenges are already reported in the literature. Still, despite the fact that the majority of the articles in this SLR address issues related to user interaction with large volumes of data, only the eight most relevant studies—according to the applied quality criteria—were selected for highlighting. Problems related to data visualization, as illustrated in Figure 4, are mentioned by authors such as Ferreira et al. (2024), Tylosky et al. (2025), Barcellos et al. (2018, 2022), which shows that these are recurring and well-documented challenges that still persist in practice.

Regarding the challenges related to data, as illustrated in Figure 5, Barcellos et al. (2022, 2023, 2024) and Cantador et al. (2021) address this theme. Among the main problems identified by these studies, traceability (origin) and data reliability stand out, aspects repeatedly pointed out by all the authors.

Regarding processes and interaction, these aspects were addressed by Xiao et al. (2023), Cantador et al. (2021), and Barcellos et al. (2022). Ferreira et al. (2024), Tylosky et al. (2025), Barcellos et al. (2022), and Cantador et al. (2021), on the other hand, mainly focus on problems and challenges related to the context and expected user characteristics, such as the level of data literacy—even among those who are not experts.

Figure 5 - Challenges and problems identified in the literature regarding data.

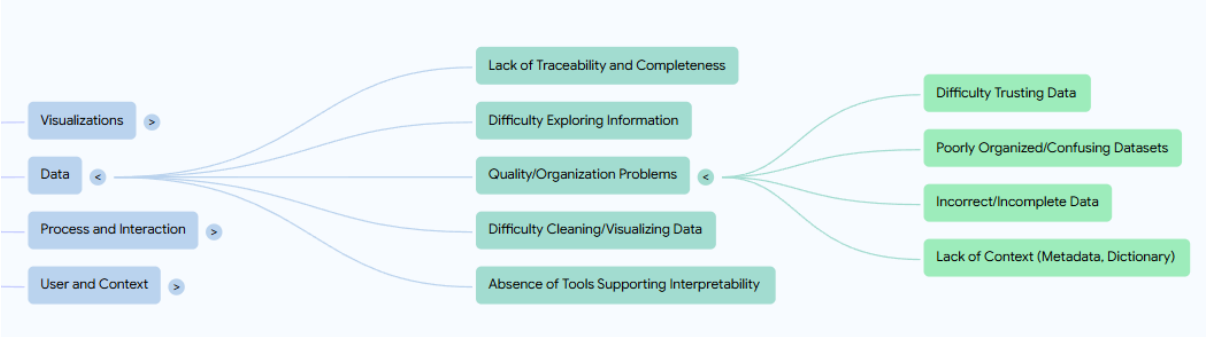


Figure 6 - Challenges and problems identified in the literature regarding interaction.

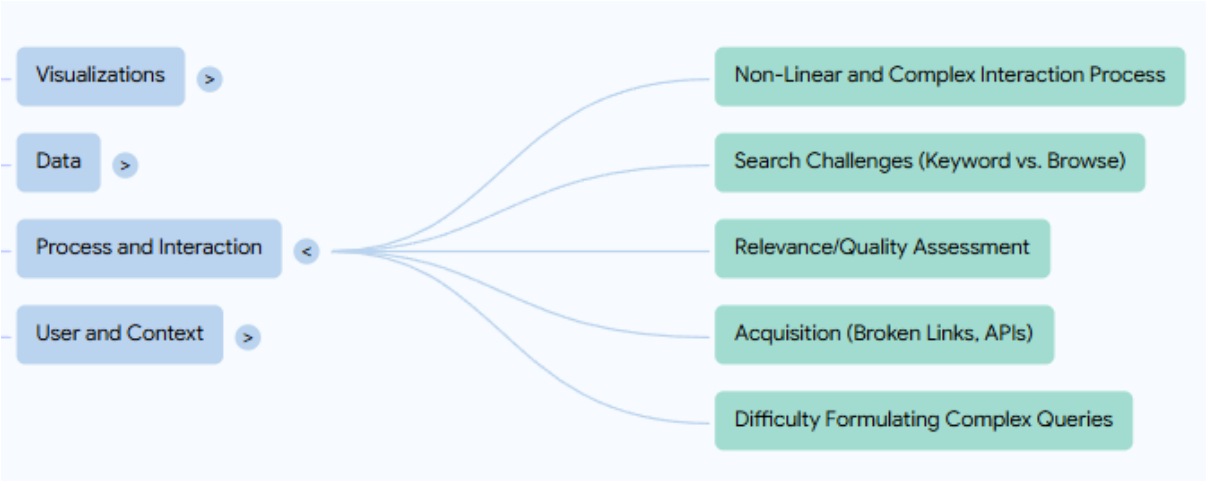
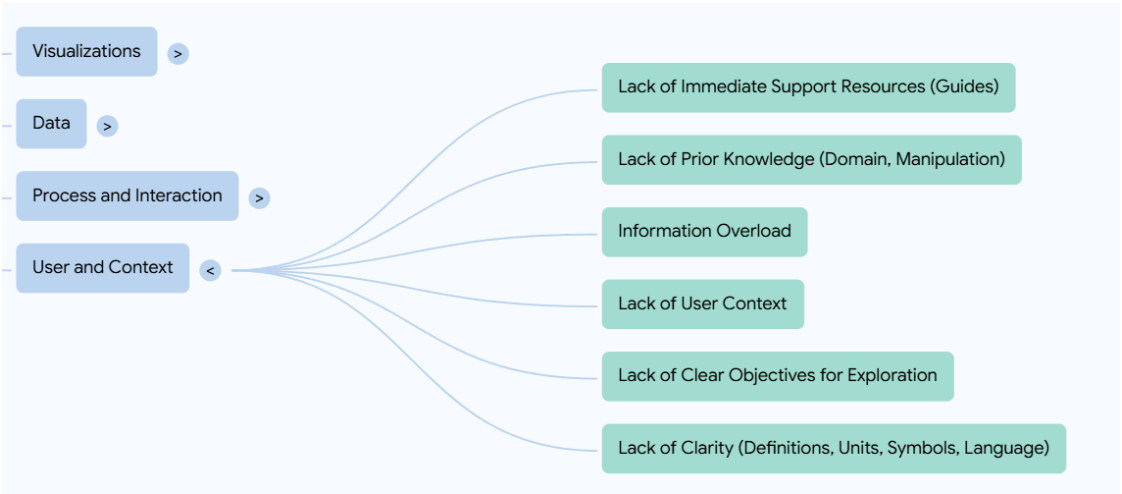


Figure 7 - Challenges and problems identified in the literature regarding context.



3.5. Discussion

Based on the results presented in section 3, challenges were identified that can serve as a basis for guiding the development of an artifact with the proposal to analyze and evaluate an open data platform, such as:

- Lack of evaluation frameworks focused on HDI: The SLR identified that most studies do not present methods, metrics, or frameworks specifically designed to evaluate open data platforms from the perspective of HDI. This justifies the need for an artifact that considers specific aspects of this interaction.
- Importance of data interpretability: Barcellos et al. (2022) highlight fundamental dimensions for interpretability, such as traceability, intelligibility, and contextualization of data.
- Need for support for non-expert users: Studies such as Tylosky et al. 2025; Ferreira et al. 2024, indicate that the absence of support mechanisms and inclusive design hinders data exploration by non-expert audiences.
- Recurring challenges in data visualization: Problems such as information overload, inappropriate use of charts, and lack of customization options were recurrent (Barcellos et al. 2018; Ferreira et al. 2024).
- Lack of well-defined interaction processes: The literature indicates that many portals lack interactive flows that guide the user in searching, filtering, and analyzing data (Xiao et al. 2023; Cantador et al. 2021).
- Absence of contextual information about the data: The lack of metadata, explanations about the origin, periodicity, and reliability of the data impairs its interpretation and use (Barcellos et al. 2023; Cantador et al. 2021).
- Insufficient integration with user-centered design practices: Few studies show evidence of the use of participatory approaches in the design of the platforms.

From the analysis of the 25 selected studies, it was possible to identify a set of challenges that can serve as a basis for the creation of evaluation criteria for the development of the artifact for analyzing and evaluating the quality of open data platforms. These findings show that, although the literature does not yet have consolidated frameworks specifically focused on HDI in open data platforms, there is sufficient theoretical and empirical input to guide the creation of a robust evaluative artifact, centered on user needs and aligned with the requirements of informational and interactional quality of these platforms.

4. Final considerations

This SLR aimed to identify frameworks, metrics, methods, methodologies, and heuristics applicable to the evaluation of open data platforms from the perspective of HDI. Additionally, the literature was investigated for interaction challenges and user needs, including non-expert users, as well as the improvements conceived within HDI to mitigate these challenges. Based on the results of the SLR, it was possible to achieve the objective of the article, that is, to identify quality criteria in open data platforms. These criteria will serve as the basis for the development of an artifact that will assist in the analysis and evaluation of the quality of this type of platform. The results obtained not only answered the research questions but also offer support for the improvement of platforms such as the ELLAS Platform (access here), and directly align with one of the Grand Challenges of HCI [Pereira et al. 2024] (GranDIHC-BR 2025-2035 - GC5: Human-Data Interaction, Data Literacy, and Usable Privacy [Coleti et al. 2024]).

Regarding the definition of HDI (RQ1), the SLR revealed a consistent understanding across the analyzed studies, characterizing it as an interdisciplinary field that goes beyond traditional usability. HDI encompasses the ability to manipulate and analyze data, support decision-making, personalization, understanding of behavioral impacts, and the principles of Legibility, Agency, and Negotiability (Mortier et al., 2014; Barcellos et al., 2022). For the ELLAS platform, the internalization of these concepts is

imperative. It is worth noting, however, a methodological limitation of the review: although the topic of HDI has been discussed since 2010, the inclusion criteria adopted restricted the analysis to publications between 2020 and 2025, which may have led to the exclusion of relevant studies published earlier and not found in the automated search, as well as in the manual search (Snowballing).

An important result of this SLR (evidenced in the answer to RQ2 and analysis of QQ1/QQ2) is the prominent gap in the existence of frameworks, models, methods, metrics, or heuristics specifically designed to evaluate the quality of open data platforms from the integral perspective of HDI.

The challenges and needs of users (RQ3), mapped in Figures 4 to 7, list problems identified in Visualization (e.g., difficulty in differentiating types, mapping of variables, problems for colorblind users), Data (e.g., lack of traceability, difficulty in trusting the data, absence of metadata), Interaction (e.g., non-linear process, challenges in searching, difficulty in formulating queries), and Context (e.g., lack of guides, information overload, lack of clarity in definitions) are precisely the obstacles that the ELLAS Project must seek to mitigate or eliminate. The ELLAS Platform's interface should be intuitive, facilitating exploration and discovery, even for non-expert users, which connects to the "Data Literacy" advocated by GC5. The availability of support resources, such as tutorials and contextual explanations, is equally crucial.

During the analysis of the literature, the need for a more precise definition of the concept of "quality evaluation" in open data platforms, from the perspective of HDI, became evident for future investigations and for the structuring of the evaluation framework to be developed based on this review. Thus, the findings of this SLR not only fulfill their objectives but also provide input for the future evaluation of the ELLAS platform. By addressing the identified gaps and challenges, the ELLAS project and the associated research can contribute remarkably to the advancement of the field of HDI, aligning with the goals of GranDIHC-BR GC5 to promote a more effective, empowering, and inclusive HDI, especially for non-expert users.

Given the identified gaps and challenges, this study points to opportunities for future research, such as the need to develop and validate frameworks, models, methods, and tools to evaluate the quality of open data portals, focusing on the dimensions of HDI. These frameworks should go beyond technical evaluation and consider aspects such as interpretability, accessibility, usability, and user experience, including the perspective of non-expert users. Furthermore, future research can explore innovative solutions to mitigate the interaction challenges identified in this review. This effort is in line with GC5, which points to HDI as a field in its initial stages that demands research and development, including forms of evaluation, development of tools and techniques to improve data interpretability, provide contextualized support, facilitate manipulation and analysis by non-expert users, and promote accessibility in open data portals.

5. Acknowledgement

To the ELLAS project, for the opportunity to use the ELLAS Connected Open Data Platform as a case study, to CAPES (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*), to CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) and *Pró-Reitoria de Pesquisa* from UFMT for partially funding this research. We also thank the use of the artificial intelligence tools ChatGPT and Gemini, employed in the review and textual improvement process. We highlight the use of the

NotebookLM tool, which contributed to the systematization, analysis, and study of the articles selected in the SLR.

6. Ethical Precautions

The ethical considerations for this research have been carefully addressed. As this study is a systematic literature review, it is based entirely on the analysis of previously published and publicly accessible scholarly literature. No human participants were directly involved in the research process, and no new data, personal information, or images of individuals were collected or processed for this study. In accordance with institutional and international guidelines, research that does not involve human subjects is exempt from review by an ethical committee.[2 All sources used in this review have been appropriately cited to respect intellectual property and ensure academic integrity.

7. References

Barcellos, R., Bernardini, F. and Viterbo, J. (2022). “Towards defining data interpretability in open data portals: Challenges and research opportunities.” In Information Systems, v. 106, p. 101961.

Barcellos, R., Bernardini, F., Viterbo, J. and Zuiderwijk, A. (2023). “Hippolyta: a framework to enhance open data interpretability and empower citizens.” In Proceedings of the 24th Annual International Conference on Digital Government Research, pp.191–198, ACM.

Barcellos, R., Bernardini, F., Zuiderwijk, A. and Viterbo, J. (2024). “Exploring Interpretability in Open Government Data with ChatGPT.” In Proceedings of the 25th Annual International Conference on Digital Government Research, (pp. 186-195).

Barcellos, R., Viterbo, J., Bernardini, F. and Trevisan, D. (2018). “An Instrument for Evaluating the Quality of Data Visualizations”. In 22nd International Conference Information Visualisation. (IV) (pp. 169-174). IEEE.

Cantador, I., Viejo-Tardío, J., Cortés-Cediel, M. and Pedro, M. (2021). “A Chatbot for Searching and Exploring Open Data: Implementation and Evaluation in E-Government.” In Proceedings of the 22nd Annual International Conference on Digital Government Research. (pp. 168-179).

Coleti, T. A., Zacarias, R., Divino, S., Saraiva, J., Salgado, A., Gonçalves, D., Morandini, M., Santos, R. (2024) “GranDIHC-BR 2025-2035 - GC5 - Human-Data Interaction, Data Literacy and Usable Privacy” In Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems. (pp. 1-24).

Degbelo, A. (2020). “Open Data User Needs: A Preliminary Synthesis.” In Proceedings of the Web Conference 2020, pp.834–839.

Elmqvist, N. (2011). “Embodied Human-Data Interaction” In ACM Workshop “Embodied Interaction: Theory and Practice in HCI , (Vol. 1, pp. 104-107).

Ferreira, B., Rodrigues, A., Barbosa, F., Barbosa, S. (2024). “Investigating Difficulties and Decisions in Information Visualization Design” In IHC '24: Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems (pp. 1-11).

Flausino, M., Kozievitch, N., Fonseca, K. and Liu, E. (2024). “Open Data Portals - A case study of Challenges and Opportunities.” In Brazilian e-Science Workshop (BreSci) (pp. 104-111). SBC.

Garcia, I., Andrade, E., Bernardini, F. and Viterbo, J. (2022). “Assessing the Quality of Covid-19 Open Data Portals.” In International Conference on Electronic Government, (pp. 212-227). Cham: Springer International Publishing.

Gil, A. C. (2008). Métodos e técnicas de pesquisa social, 6a edição. Editora Atlas.

Grant, M. J., & Booth, A. (2009). “A Typology of reviews: An Analysis of 14 Review Types and Associated Methodologies” In Health Information & Libraries Journal, 26(2), 91–108.

Hornung, H., Pereira, R., Baranauskas, M.C.C., Liu, K., 2015. “Challenges for Human-Data Interaction – A Semiotic Perspective” In 17th International Conference on Human-Computer Interaction. (pp. 37-48). Cham: Springer International Publishing.

Hsu, J., Ravichandran, R., Zhang, E. and Keung, C. (2021). “Open Data Standard and Analysis Framework: Towards Response Equity in Local Governments.” In 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, pp.1–8.

Isotani, S., and Bittencourt, I. I. (2015). “Dados abertos conectados: em busca da web do conhecimento” Novatec Editora.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. 2.

Kitchin, R and McArdle, G. (2016). “What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets” In Big Data and Society. v. 3, n. 1, p. 2053951716631130.

Koesten, L., Simperl, E., Blount, T., Kacprzak, E. and Tennison, J. (2020). “Everything you always wanted to know about a dataset: Studies in data summarisation.” In International Journal of Human-Computer Studies, v. 135, p. 102367.

Kotis, K., Vouros, G.A. (2006). “Human-centered ontology engineering: The HCOME methodology.” In Knowl Inf Syst 10, 109–131 .

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). “Benefits, Adoption Barriers and Myths of Open Data and Open Government.” In Information Systems Management, 29(4), 258–268.

Lee, B., Choe, E., Isenberg, P., Marriott, K., Stasko, J. and Rhyne, T.-M. (2020). “Reaching Broader Audiences With Data Visualization.” In IEEE Computer Graphics and Applications, 40(2), pp.82–90.

Liang, Y., Cao, Y., Chen, M., Dong, H. and Wang, H. (2025). “Determinants of open government data continuance usage and value creation: A self-regulation framework analysis.” In Government Information Quarterly, v. 42, n. 2, p. 102022.

Máchová, R. and Lnenicka, M. (2017). “Evaluating the Quality of Open Data Portals on the National Level.” In Journal of theoretical and applied electronic commerce research, v. 12, n. 1, p. 21-41.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., Crowcroft, J. (2014). “Human-Data Interaction: The Human Face of the Data-Driven Society” in SSRN Electronic Journal.

Natvig, M., Jiang, S., Stav E. (2021)”Using open data for digital innovation: Barriers for use and recommendations for publishers” In JeDEM-eJournal of eDemocracy and Open Government, 13(2), 28-57.

Oliveira, R., Cappelli, C. and Oliveira, J. (2023). “Guia de Linguagem Simples para Visualização de Dados: Construção, Design, Avaliação e Contribuições”. In Anais do II Workshop Investigações em Interação Humano-Dados, (pp. 76-82). SBC.

Pereira, R., Darin, T. and Silveira, M. (2024). “GranDIHC-BR: Grand Research Challenges in Human-Computer Interaction in Brazil for 2025-2035”. In IHC '24: Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems. (pp. 1-24).

Pinto, H.S., Tempich, C., Staab, S. (2009). “Ontology Engineering and Evolution in a Distributed World Using DILIGENT”. In: Staab, S., Studer, R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. (pp. 153-176). Berlin, Heidelberg: Springer Berlin Heidelberg.

Puussaar, A., Montague, K., Peacock, S., Nappey, T., Anderson, R., Jonczyk, J., Wright, P. and James, P. (2022). “SenseMyStreet: Sensor Commissioning Toolkit for Communities” In Proceedings of the ACM on Human-Computer Interaction, v. 6, n. CSCW2, p. 1-26.

Runeson, P., Skoglund, M. (2009). “Reference-based search strategies in systematic reviews” In Proceedings of the 13th International Conference on Empirical Assessment and Evaluation in Software Engineering. Electronic Workshops in Computing (eWIC). BCS, Durham University, UK

Sommerville, Ian (2011). *Engenharia de Software*. 9. ed. — São Paulo: Pearson Prentice Hall

Tylosky, N., Knutas, A., Qureshi, M. and Wolff, A. (2023). “User Experience for Non-Expert Audiences in Data Exploration.” In Proceedings of the 11th International Conference on Communities and Technologies. p. 12-16.

Tylosky, N., Knutas, A. and Wolff, A. (2023). “Principles for Designing for Data Exploration for a Non-Expert Audience.” In 46th ICT and Electronics Convention - Proceedings, pp.88–91.

Tylosky, N., Knutas, A. and Wolff, A. (2025). “Design practices in visualization driven data exploration for non-expert audiences”. In Computer Science Review. v. 56, p. 100731.

Vasconcelos, L., Barcellos, R., Viterbo, J., Bernardini, F. and Salgado, L. (2020). "Investigating Communicability Issues in the Open Data Manipulation Flow." In AMCIS 2020 Proceedings.

Victorelli, E. and Reis, J.C. (2020). "Human-data interaction design guidelines for visualization systems." In Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems,pp.1–10.

Victorelli, E., dos Reis, J., Hornung, H., Prado, A. (2019). "Understanding Human-Data Interaction: Literature Review and Recommendations for Design" In International Journal of Human-Computer Studies. v. 134, p. 13-32.

Xiao, F., Chi, Y. and He, D. (2023). "Promoting data use through understanding user behaviors: A model for human open government data interaction." In Journal of the Association for Information Science and Technology, v. 74, n. 13, p. 1498-1514

Xiao, F., Thaker, K. and He, D. (2022). "Categorizing Open Government Data Users by Exploring their Challenges and Proficiency." In CHI Conference on Human Factors in Computing Systems, pp.1–7.

Xiao, F., Wang, Z. and He, D. (2020). "Understanding users' accessing behaviors to local Open Government Data via transaction log analysis." In Proceedings of the 83rd Annual Association for Information Science and Technology. v. 57, n. 1, p. e27