Feature Article

PhDs – Portugal has Doctors: A Visualization of Academia Achievements in Portugal from 1970 to 2022

Pedro Silva, *University of Coimbra, CISUC, DEI*

Pedro Martins, *University of Coimbra, CISUC, DEI*

Penousal Machado, *University of Coimbra, CISUC, DEI*

Abstract—PhDs – Portugal has Doctors is an interactive installation presenting Portuguese doctoral theses from 1970 to 2022, tracking their historical evolution and distribution across universities and research sectors. This work resulted in an installation that served a dual purpose: to raise awareness and value the work of national doctorates and to reduce the communication gap on this topic, encouraging a public engagement with the subject, fostering discussions beyond the data and prompting reflection on how this lesser-known reality has impacted Portugal with its significantly growing presence. Drawing from recent research on aesthetics and its impact on data perception, we integrated these insights to make the visualization approach accessible to the general public, emphasizing a visually minimalist narrative. The story initially prompts viewers to reflect on the temporal evolution of the academic landscape over the last decades. The visualization then encourages viewers to actively engage with the data, facilitating a more in-depth exploration. The installation garnered positive feedback, provoking amazement and surprise, and revealed an unknown reality, even within the scientific community further justifying the need for this type of dissemination works.

In contemporary society, the academic community, particularly doctoral researchers, plays a crucial role in generating knowledge, driving innovation, and shaping the trajectory of various sectors. However, despite their significant contributions, doctoral graduates often face challenges in accessing opportunities and achieving recognition within their respective fields. Moreover, their research findings and the broader academic landscape are not always effectively communicated to the general public, resulting in a lack of awareness regarding the significance of doctoral research and its societal implications. As members of the PhD community ourselves, we are deeply invested in showcasing the academic value of doctoral research and enhancing its visibility in society. This personal motivation highlights the importance of initiatives aimed at exploring and disseminating information about doctoral research, which can contribute to a deeper understanding of academic contributions, their relevance to society and, ultimately, lead to sociopolitical measures and efforts to improve academic professionals conditions with more fundings, scholarships and employment opportunities from both public and private entities.

Various studies have highlighted the crucial role of doctoral graduates in generating knowledge within universities and transferring it to companies and organizations. They are essential for driving innovation across different industries and the knowledge economy [1]. However, in some countries, these graduates do not have satisfactory working conditions. In Portugal, for instance, factors such as salary, stability, and recognition have led many young and highly educated professionals to seek opportunities abroad [2]. Several studies have examined the challenges faced by doctoral graduates in Portugal, emphasizing the urgent

XXXX-XXX © 2021 IEEE
Digital Object Identifier 10.1109/XXX.0000.0000000

need for policy interventions to leverage their skills and improve the country's economic competitiveness and scientific ecosystem [3]. In the realm of Portuguese arts, initiatives like DesignOBS aim to promote design as a driver of innovation, cultural development and socioeconomic growth. However, design doctorates often lack visibility, and doctoral research in design faces challenges such as limited continuity and reproducibility [3], [4]. This is unfortunate because designers and artists have an incisive power to unearth and bring to light themes and information in rather provocative ways that may be overlooked or misrepresented by social and political entities. Similarly, our work aims to shed light on national doctorates, exploring their numbers, historical trends, and distribution across universities and research sectors. We seek to engage the public with data in public settings, encouraging reflection, discussion, and a deeper understanding of research findings [5]. From the public perspective, interest in scientific research dissemination, particularly from doctoral studies, can span across different audience groups with distinct concerns and motivations. For the general public, here understood as every person outside the academic and professional context, such data offers insights into groundbreaking discoveries and innovations that can impact industries, shape policies, and ultimately enhance daily life. Within the PhD community, there is a growing effort to make research data and information more accessible and engaging to the public so that everyone can understand their work, and potentially promote their field and research to attract more practitioners or investors. In contexts like public exhibitions, interactive visualizations can also operate as educational catalysts, captivating audiences to explore and learn more about the presented data.

Moreover, we strive to improve the accessibility and dissemination of these datasets and associated information, which are often underutilized due to limited communication, educational, or technical barriers [6]. This challenge arises from the limited communication provided by their repositories and websites, coupled with the restricted possibilities they offer to retrieve and manipulate data, often making it difficult to analyze since users lack the skills to access and work with this data effectively.

THE ARTIST ROLE

Over the centuries, professionals from other fields than Information Visualization have made relevant contributions with good examples of how to use visualization to communicate a message in a persuasive way to the general, non-scientific public. Some pivotal references on this note are Florence Nightingale's rose diagrams to expose the poor sanitary conditions as the main reason for Crimean War deaths, Joseph Minard's graph to depict the consequences of Napoleon's March to Moscow, John Snow's map plotting the cases of cholera to find the outbreak origin, and Du Bois' visualizations depicting a series of statistics illustrating the live condition of the Black Americans to show the progress made by freed slaves and their descendants post-emancipation. Nowadays, with easier access to datasets and technology, artists have played a significant role in popularizing data visualization at the turn of the century, with some noteworthy visualizations by authors such as Brendan Dawes, Ben Fry, Golan Levin, Fernanda Viégas, Jason Salavon, among many others. More recently, other communication professionals, from illustrators to journalists such as Federica Fragapane, Mona Chalabi, Nadieh Bremer, or Shirley Wu, have also made important strides in introducing data visualization to broader audiences through compelling approaches such as infographics or anthropographics [7], [8] due to their ability to make them connect more intimately with data. This communication effort has also been visible in academia, where scientific spaces such as IEEE VISAP and the SIGGRAPH Arts program have created space and fostered the creation of works that merge visualization with the arts.

These emerging professionals and artists blend techniques from diverse sources, resulting in unconventional outputs that offer unique insights into data and the scientific field's usefulness [9]. Unlike traditional scientific communication, artists aim to make complex ideas understandable to a wider audience, intending to provoke new thoughts and emotions or raise awareness about underlying data issues. They achieve this through many different approaches, ranging from compelling installations and sculptures [10] to inventive visual metaphors and embellishments [11], rather than presenting data in conventional or prosaic ways. Artists started to promote a new attitude, "embracing the fact that visualizations can be used to persuade as well as analyze" [12]. With an increasing number of practitioners prioritizing user experience [13], artists are also striving to emotionally connect people with important topics, fostering deeper understanding and awareness about a given subject. However, reaching broader audiences poses both challenges and opportunities.

These wider audiences may have diverse needs and contexts for using data, which can impact how and why data is accessed. They might not be interested in in-depth data exploration, using data visualization more casually for various purposes and in different setups or

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3405656

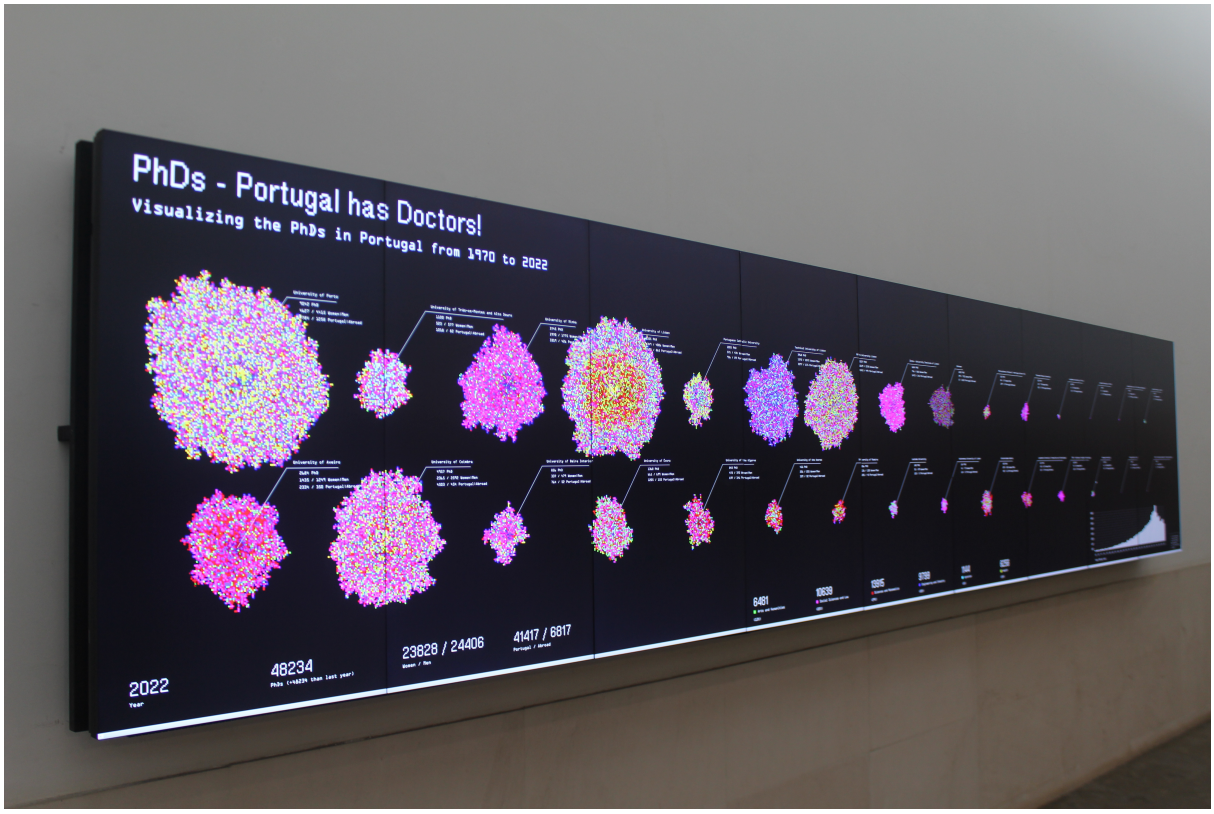

FIGURE 1. Photograph of the video wall containing the initial page of the visualization, where the entire dataset is depicted with the doctorates distributed across the universities where they were carried out.

displays [14]. Therefore, the role of artists, especially in public art projects, could be a solid force to bridge the gap between scientific research and the public by creating works that engage, inspire, and communicate scientific subjects in a more appealing way. This not only promotes greater awareness and understanding of important issues, but it also highlights the relevance of art and its ability to foster social and cultural change [15]. In our project, we took an artistic approach to visualize our dataset, leveraging its benefits to create a comprehensive visualization that rewards aesthetics and vernacular communication as a gateway to user engagement with the data that can ultimately lead to deeper analytical explorations once the user is already actively engaged and immersed in the visualization.

THE INSTALLATION

Manovich introduced the concept of *Cultural Analytics*, a study involving the analysis of large cultural datasets using computational and visualization techniques [16]. Building upon this approach, we incorporate artistic and aesthetic elements into our visualization as further research has shown that aesthetics can enhance data perception and engagement, suggesting that it can improve our understanding of our society and our everyday lives [17]–[20].

MOTIVATION

PhDs – Portugal has Doctors is a web-based interactive installation that depicts every doctoral thesis from 1970 to 2022 either done in Portugal, at a Portuguese university, or conducted abroad with further equivalence attributed by a Portuguese higher education institution. The work establishes itself as a data visualization work operating as a creative artistic practice with which we want to highlight what seems to be an undervalued or even forgotten number of achievements in national research and development. To achieve this, we endeavor to create a large-scale portrait that assumes the role of a narrative, a visual story that brings forth the essence of our academic community. This is where the practice of traditional data visualization transcends its mere data communication boundaries and takes on the qualities of a

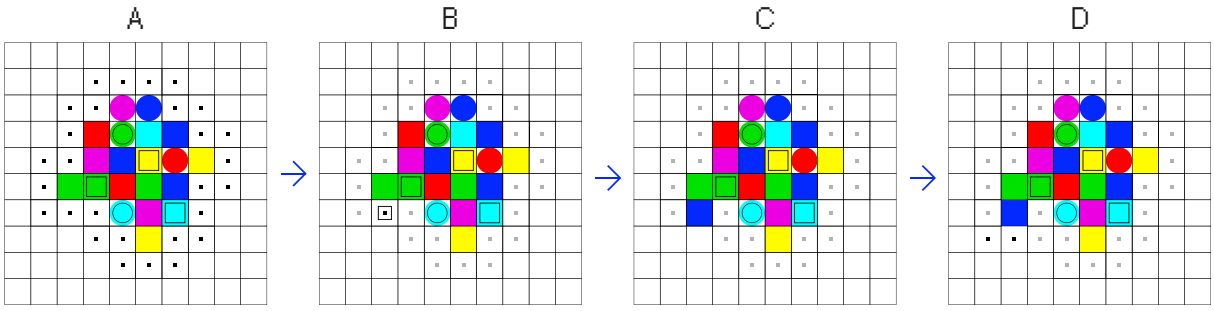

FIGURE 2. Each university serves as a starting point, with adjacent empty cells identified (represented by small black dots) and placed in a list (A). The next thesis position is randomly selected from the list of options (B). The position is filled with the new thesis element (C), and the list is updated by removing that position and retrieving the new available positions (D). The process repeats until there are no more theses to be positioned.

creative artistic endeavor. It becomes more than just a means of representing data, it evolves into a form of expression, a medium through which we not only convey information but also engage the viewer on an emotional and societal level.

THE DATASET

The data used in this work is available on the RENATES platform at https://renates2.dgeec.mec.pt/, which results from an official information collection by Portuguese higher education institutions and the General Directorate of Higher Education. Since the database limits the consultation and downloading of its data, we made a one-off request to RENATES to supply all the information available at the time (28 December 2022). The dataset provided contains 48,234 entries. Each entry represents a doctorate, containing the author's name, gender (already labeled with only feminine/masculine identifications), thesis title, country and university where it was granted, course name, scientific domain, and year of completion. The dataset includes doctorates associated with 30 universities, including both universities from mainland Portugal and the autonomous regions of Madeira and the Azores. It covers 42 academic courses that are encapsulated in 6 distinct scientific domains: (i) Arts and Humanities; (ii) Social Sciences and Law; (iii) Science and Mathematics; (iv) Engineering and Industry; (v) Agrarian; and (vi) Health. These six field of study classifications align with the CNAEF - National Classification of Education and Training Areas, which categorizes courses and study programs offered by higher education institutions. The dataset and the project code are available at https://github.com/PedroNunoSilva22/PhDs.

THE VISUALIZATION

The entire dataset was accommodated on a single display, on a video wall consisting of 9 LCDs with a resolution of 9720 x 1920 pixels (Fig. 1) available in our department. We imagined the visualization as a grid of cells, each representing a potential position to a single doctoral thesis.

In a very simple way, a doctoral thesis involves creating new knowledge, supported by scientific publications validated by peers and/or by obtaining patents. Creating new contributions results in an incremental evolution of knowledge, building on top of already consolidated prior knowledge. To illustrate how knowledge grows with each new doctoral thesis, we developed an algorithm inspired by the Breadth-First Search Algorithm [21]. The Breadth-First Search Algorithm was chosen as a reference for our visual representation due to its ability to systematically explore and traverse a graph or grid in a structured manner, facilitating the placement of new theses within the visualization grid in order to visually represent the incremental contribution of knowledge with each new thesis in a logical and coherent manner. However, we introduced a twist in the algorithm: rather than following a predetermined ordered list of options, our algorithm randomly selects the next position to place a doctoral thesis. This deliberate randomness was employed to prevent the visual output of a predictable continuous spiral, where each new thesis would appear directly linked and placed next to its predecessor, which could suggest non-existent continuity relationships. By incorporating randomness, we aimed to reflect the diverse and most of the times unrelated nature of research topics and theses. This approach ensures that the visual representation captures the organic growth of knowledge, where areas of exploration can vary widely and unexpected pathways

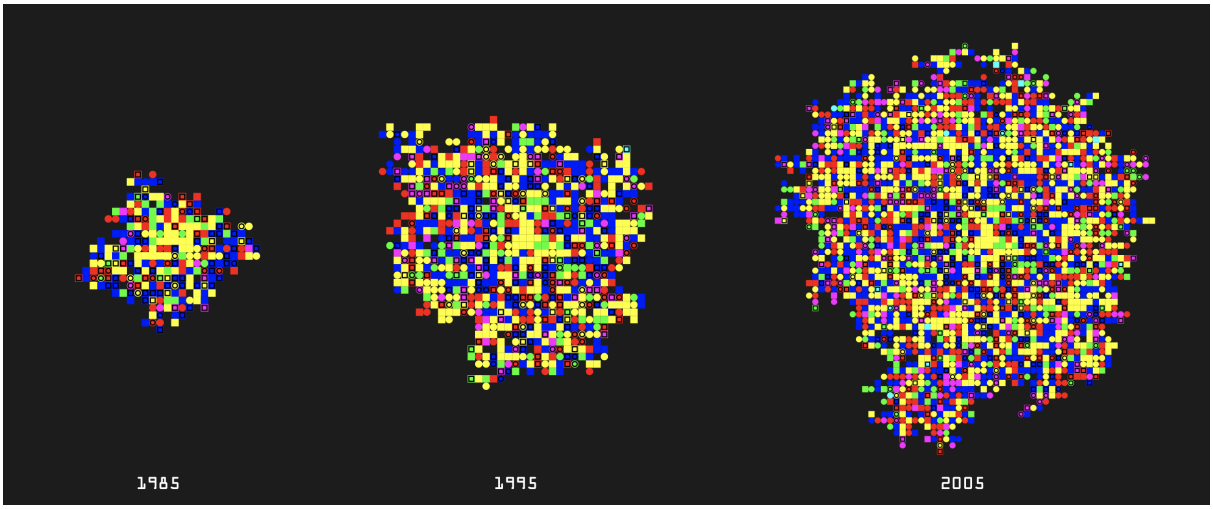

FIGURE 3. Snapshots of the visual evolution of the doctorates held at the University of Porto from 1985 to 2005, depicting the contributions in 1985, 1995, and 2005.

may emerge. The visual representation resulting from our approach can be seen in Fig. 2. Every university in our dataset has a starting point in our visualization grid, from which we start to place the corresponding theses. As each thesis is added, we identify adjacent empty cells (represented by small black dots), adding them to a list of potential positions for the next thesis. The next position is chosen randomly from this list, and the process repeats until all theses have been placed. Because of the random selection, gaps may appear in the grid, resulting in paths of varying lengths. This mirrors the ebb and flow of research in different scientific domains over time. Some subjects may be trending and heavily researched, while others remain less explored, creating "holes" or paths waiting to be discovered.

In our visual representation, each doctorate is depicted using one of two simple visual elements: squares for male doctoral candidates and circles for female candidates. This decision was informed by research suggesting that certain basic shapes are associated with gender categories [22]. We used color to represent the areas of study present in our dataset, where we deliberately tried to select color schemes that could correspond to the color typically used in Portuguese universities to distinguish the different scientific domains. As such, (i) green was used for Arts and Humanities, (ii) pink for Social Sciences and Law, (iii) red for Science and Mathematics, (iv) blue for Engineering and Industry, (v) cyan for Agrarian and (vi) yellow for Health. Furthermore, the decision to use a subtle black inner outline for doctorates completed partially or entirely abroad serves a dual purpose. Beyond simply indicating this information visually, the stroke operates metaphorically as a "border". This choice symbolizes the crossing of borders and the international nature of these doctorates (Fig. 2).

The initial page of the visualization shows us the doctorates aggregated by the university in which they were carried out (Fig. 1). This way, we have a complete view of the national panorama across the years depicted in a single display. By default, this initial page iterates over the years to reveal the evolution and distribution of doctoral contributions across the country. As more new PhDs are added each year, the forms of each university become larger and larger (Fig. 3).

In addition to the forms of universities generated by their corresponding doctoral theses, our visualization incorporates a selection of statistical data points. These statistics encompass various dimensions of the doctoral landscape, including the total number of doctorates completed to date, the distribution of doctorates across different academic domains, the breakdown of doctorates by gender, and the geographical locations where these doctorates were completed. At first glance, the juxtaposition of artistic visualizations with statistical data may appear contradictory, given the initial framing around the focus of such visualizations on aesthetics, rather than readability or analytical accuracy. Yet such combination encourages a more profound engagement where the initial aesthetic experience leads to an analytical exploration of the data. While the artistic visualizations serve as an entry point for user engagement and contemplation, drawing

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3405656

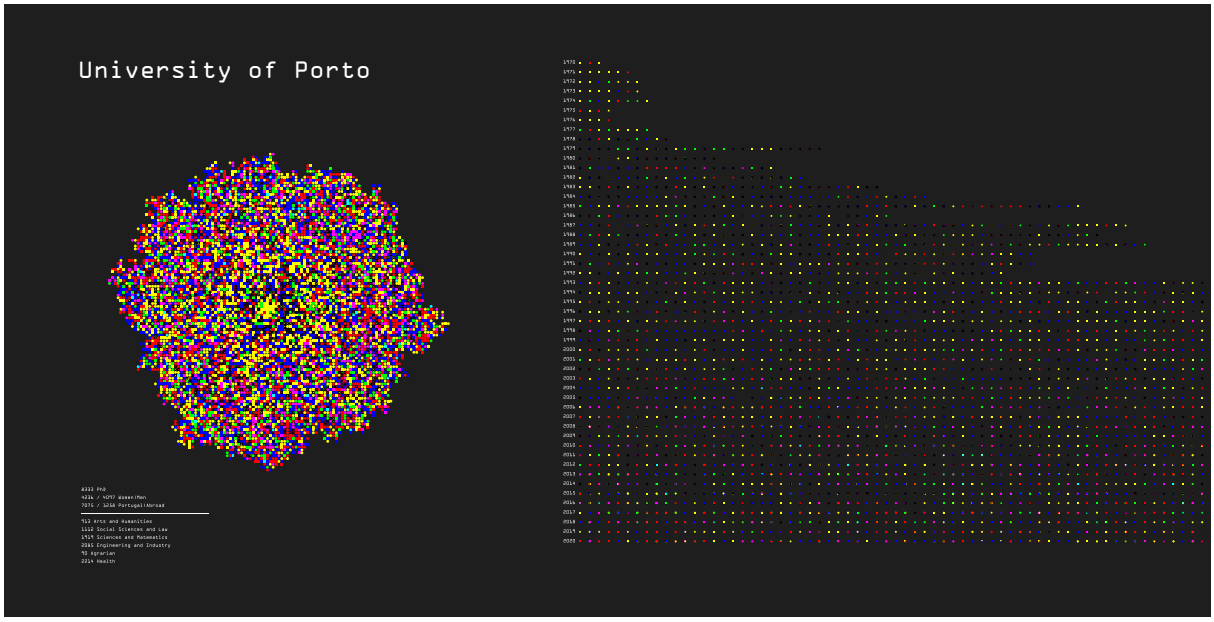

FIGURE 4. Snapshot of the single university view. In this view, on the left, we have a close-up of the university cluster of doctorates with some statistics and, on the right, we have these same doctorates aligned vertically and organized by year.

attention to the broader narrative of academic accomplishment, the inclusion of statistical data provides a pathway to deeper exploration and interpretation.

In this layered approach, the viewer first encounters global statistics that offer an overview of the dataset, setting a contextual foundation. Viewers interested in deeper insights can interact with the visualization (explained later on) to explore more specific information on another page, the university page which includes extra details about a given university and its associated doctoral theses.

When on the university page, the user can individually consult the details about a specific doctoral thesis. By default, the university's first PhD is visualized. When selecting the desired doctorate, its details (author's name, thesis title, area of study, and year of completion) are presented accompanied by an abstract representation generated with these same details. This abstract representation tries to unfold the approach used in the initial page of the visualization to represent doctoral theses with just one element (square or circle). Here, that single element is broken down into multiple components to represent its multiple associated details. The name of the doctorate's author is represented either by purple squares, if masculine, or by pink circles, if feminine. Outlined squares are used to represent the doctorate's field of study and colored with the same color coding as previously enumerated apart from the social sciences and law that take an orange color instead of the original pink to distinguish themselves from the feminine pink color. The title of the thesis is illustrated with black glyphs and it is where we can distinguish national from abroad doctorates. If done abroad, the glyph takes the same black outline used in the initial visualization. Lastly, the year is represented with black crosses (Fig. 5).

This algorithmic approach was chosen to represent each of the doctorate details mentioned above because it maintains consistency with the initial grid-filling method used to map the PhDs in the visualization's initial page. However, we modified the algorithm to select the next cell unit based on different criteria. Now, each cell unit is used to represent a character of a certain detail (e.g. name), and the direction taken to place the next element in the grid is done according to the character octal value. We chose to use octal values because they provide a standardized numerical representation with the same number of outputs as the number of possible directional movements on our grid, simplifying the mapping process for determining the positioning of each detail element. Take the example depicted in Fig. 5 for the name "John Doe", where the black square represents the starting point for our algorithm. In (Fig. 5, A) we can see the eight possible directions to advance in our algorithm. These eight directions were labeled according to (Fig. 5, B) which are responsible for dictating the direction to take in order to place the next character in the grid. Since we have

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3405656

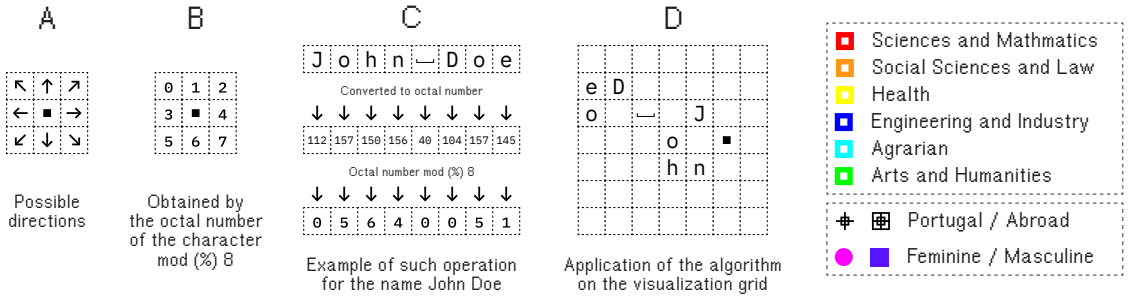

FIGURE 5. Illustration of the algorithm for representing individual doctoral details. (A) Displays potential movement directions labeled with corresponding numbers as depicted in (B). (C) Shows an example of text converted to octal values, then mapped to the corresponding number responsible for dictating the direction. Finally, (D) presents the visual application of the algorithm in the visualization grid. On the right, an overview of the visual components used to construct the different doctoral details.

eight possible directions, we chose to use the octal value of the characters to dictate such a rule. To do so, for each character of the name (spaces included), we use its corresponding octal number and then further apply the modulus operator (%) of 8 (Fig. 5, C). With this operation we obtain a number from 0 to 8, thus dictating the grid cell to proceed according to the labeling in (Fig. 5, B). Should be noted that there are no overlaps within the mapping process of a given detail, that is, if by chance the cell in the direction to be taken is occupied, the algorithm continues to proceed in that given direction until it finds an available cell. However, this overlap can occur between details because this process is independently applied to every doctorate detail (author name, thesis title, field of study, and year of completion), always starting from the same initial position. The visual application of the algorithm on the grid can be seen in (Fig. 5, D) and the corresponding visual elements of each doctorate detail are depicted in (Fig. 5, right).

Looking at Fig. 6) we can distinguish masculine from feminine and Portugal from abroad made doctorates. Unfolding each university to reveal its individual doctoral theses was a deliberate choice to showcase the extensive body of work accomplished throughout the years. Allow such detailed examination of each one of these works with such emphasis underscores our commitment to providing due and deserved recognition to each doctoral thesis within the visualization.

This design strategy containing these two distinct views with different purposes ensures that the visualization is not only a passive exhibit but an interactive experience that caters to varying levels of inquiry, from casual observation to in-depth research. By catering to both, the visualization surpasses the existing platform (RENATES) function operating as a mere repository of information and becomes a dynamic tool where the users can quickly obtain an idea of the distribution of doctorates throughout the country and time. All this in a visual way, with ease between searches in contrast to the tabular and merely textual information obtained through weary form queries on the RENATES portal.

INTERACTION

When desired, the public can engage with the installation using a wireless keyboard with an integrated trackpad, ensuring a comfortable and familiar interaction similar to that of a desktop application on a personal computer. The participants are provided with a set of hotkeys to interact with the visualization, minimizing the need for extensive scrolling giving the display dimensions. The list of hotkeys is displayed in the bottom-right corner of the screen, containing shortcuts to view just one of the six scientific domains, hiding all the others, a toggle to pause or resume the visualization loop between years, to allow the user to stop and consult the visualization data only until the desired year. As mentioned before, the audience has at her disposal some mechanisms to filter the visible doctorates, being able to drill down the dataset by navigating between the initial page and the university page. When on the initial page, by clicking on the university visual area, the correspondent university page is presented to the user, where a close-up of the selected university visual form is presented along with some extra statistics. The user can also consult all doctorates from that university, which are broken down into a grid, organized vertically by year. (Fig. 4). On this page, the user can individually select a doctoral thesis using the directional arrows on the keyboard, updating the abstract representation on the right most part of

the page.

Furthermore, the audience has the ability to group doctoral theses according to their properties. For instance, doctoral theses can be grouped by author's gender, scientific domain, location (Portugal or abroad), year, and even nested groupings of these same properties, enabling the users to discern additional insights and relationships within the dataset. In Fig. 7) we see the doctorates grouped by author's gender followed by location (national or abroad). To do so, the user must first select the grouping properties from a dropdown list in the desired order before applying the grouping process, which is also activated by a hotkey.

DISCUSSION

Our work wanted to examine the evolution of doctoral education over time, focusing on the growth in the number of PhDs awarded and the demographic distribution of these degrees. As a first personal analysis, we examined the dataset to identify trends and changes in the doctoral panorama. The dataset reveals a significant increase in the number of PhDs completed from 1970, with only 60 degrees, to a peak in 2015 with 3,163 degrees. However, there was a notable decrease of around 50% to 1,514 PhDs by 2022. This fluctuation may reflect various factors influencing doctoral education, including economic conditions, policy changes, and academic trends as previously stated. By 2022, the gender distribution of PhD holders was 23,828 women to 24,406 men, indicating a relatively balanced representation. The majority of PhDs were awarded to domestic students (41,417), compared to international students (6,817), highlighting a prevalence of national doctorates, perhaps indicating the lack of attractive international programs or financial support from academic entities or scholarships.

The dataset reveals a significant concentration of doctorates awarded by the country's largest and oldest universities, which are situated in the major cities. Specifically, the University of Porto (9042), the combined universities of Lisbon (18974), and the University of Coimbra together account for 32,973 of the 48,234 PhDs, amounting to just over two-thirds of the total.

While these institutions' prestige and historical significance contribute to their appeal, the disparity in doctoral distribution suggests that newer universities may be at a disadvantage. This imbalance could be attributed to various factors, such as the established reputation, resources, and academic networks of the older universities, which naturally attract a larger academic community. Nevertheless, it may be beneficial to explore measures that enhance the attractiveness of the newer universities. Such measures could include increasing funding for research, improving facilities, offering competitive scholarships, and promoting innovative programs that align with current and future academic and industry demands. Moreover, fostering collaborations between universities and creating incentives for faculty and students to engage with emerging institutions could help redistribute the academic population more evenly. Similarly, the disparity between the numbers of doctoral degrees awarded in different research areas suggests a potential need to explore strategies to reinforce engagement in fields with lower representation. Science and Mathematics led with 13,915 PhDs, followed by Social Sciences and Law with 10,639, and Engineering and Industry with 9,799. Arts and Humanities contributed 6,481 PhDs, Health Sciences accounted for 6,256, and Agrarian studies had the fewest with 1,144. This variance underscores the importance of considering initiatives to support and encourage doctoral research across the entire country and disciplines.

Then, we conducted an user research through informal interviews with the participants, in an attempt to avoid influencing their real-world experience, to examine how they interacted with the visualization artifact, and to identify any possible areas for improvement. More specifically, we wanted to see how people understood the visual representations and to assess if it raised awareness about the subject. The entire department community, including students, teaching, and non-teaching staff, was invited to explore the visualization over a week, yet its access was not limited to this audience since the department is open to the public. However, as would be expected, the installation was visited and tried mostly by professors, students, and other members of the department staff but we had some visitors from outside the academic community (Fig. 8).

The approaches to participants sought to be as least intrusive as possible so that, although informal, they were applied in real contexts and circumstances, where normally no one is asked about what they feel about what they are analyzing, be it a painting or even a visualization. As an introductory question, the participants were always asked if they were aware of the RENATES platform and the data it made available. Doctorates who participated quickly wanted to consult their personal representation, noting they spent a lot of time appreciating the representations from a more artistic stand while consulting other representations in search of colleagues. However, doctorate or not, no one claimed to know such a platform, which reinforces the need for such a visualization project, even if only

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3405656

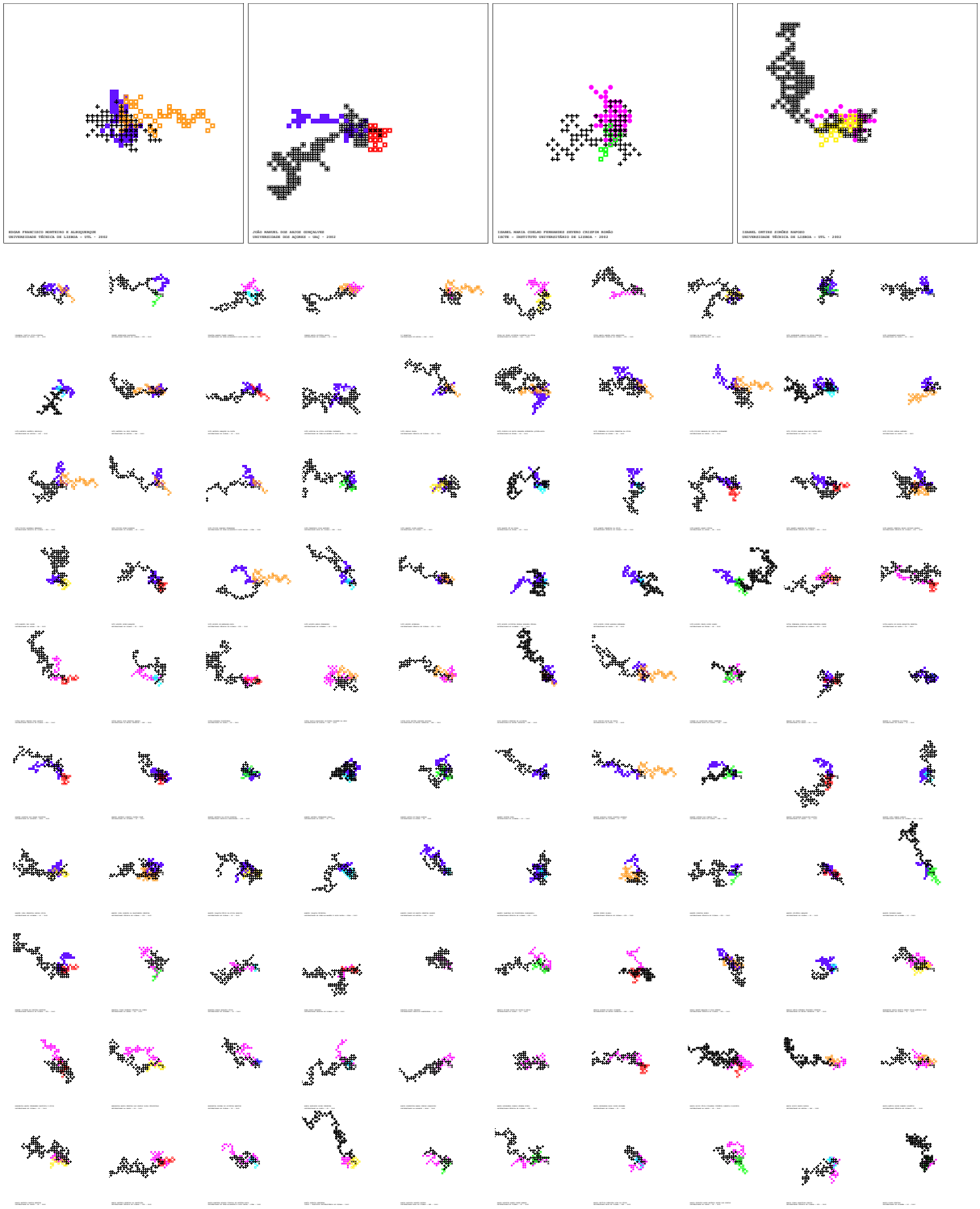

FIGURE 6. "Four PhD translated into our abstract representation algorithm revealing some possible different visual outputs with the variables used (gender, area of study, place of completion, and thesis title) (top). Collection of abstract representation showcasing one hundred (100) PhD theses (bottom)"

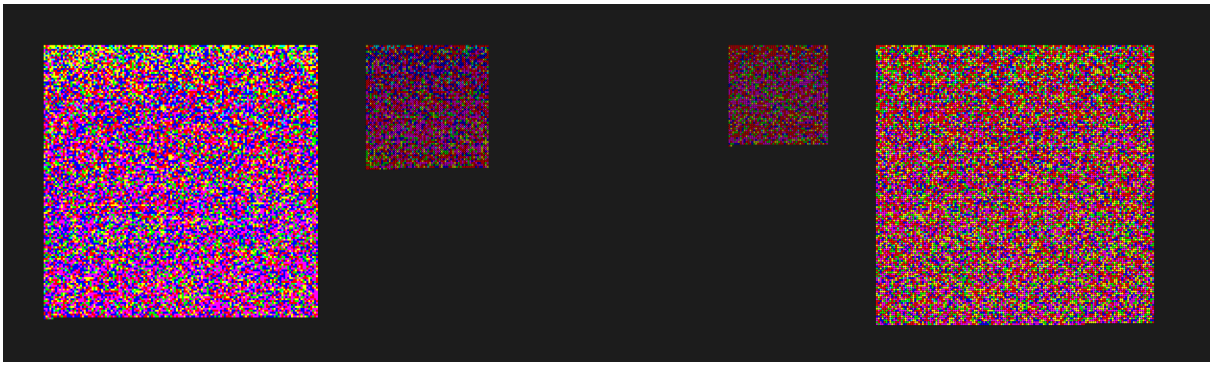

FIGURE 7. Snapshot of the visualization while grouping the entire dataset by gender and then by doctorate's local of achievement (national or abroad).

FIGURE 8. Participants interacting with the visualization.

to the extent that it helps in the dissemination of such information. Furthermore, it could lead to greater affluence and exploration of the online platform, given the interest shown by doctorates to know more about the details available about their own thesis and known colleagues or peers.

From a visual and artistic perspective, participants appreciated the overall form of the representations without delving into quantitative comparisons. Instead, they offered qualitative feedback, such as value judgments on how the author's data were visually translated and integrated into the overall composition. Initially, participants took some time to contemplate the initial page and try to make sense of what it represented without using the keyboard to interact with the visualization. Once they are acquainted with the entire information available on the page (title, statistics, hotkeys caption), they quickly discern the genre and location of execution of the PhDs, leading them to actively comment on these same variables, on how they evolved over time and how they were distributed across the different universities and scientific domains. One faculty member further noted the decrease in doctorates from 2015 onwards as something problematic, going precisely against the type of reflections and insights that we intended (Fig. 9, bottom-left). The most common observations across the public were the findings of (i) a very similar distribution between genders,(ii) the quite evident presence of a more significant number of doctorates in universities in large cities such as Lisbon and Porto, (iii) the predominance of certain scientific domains in some universities, (iv) the surprising amount of doctorates done abroad, with a ratio of 1 to 6, even in the first years recorded, and (v) the distribution of doctorates by scientific domain over the years in some universities, namely the University of Lisbon, perhaps because some faculties have been merged into the university over the years (Fig. 9, right). In general, the public seemed to have enjoyed the installation, having interacted for periods greater than 15/20 minutes, while understanding and exploring the subject portrayed, then later reflecting on the overwhelming dimension of PhDs and the numbers they were observing. This, always actively interacting with the visualization to see the temporal evolution and consulting the statistics. When exploring the visualization ourselves, we could also find large production gaps between years (Fig. 9, top-left).

Overall, most participants mentioned surprise, given the mentioned total lack of knowledge about the existence of such data, much less imagining such numbers for the subject portrayed, even for the first years of the dataset. After this surprise, they confessed to a candid engagement with visualization, with which they really enjoyed analyzing the data. This reflects the positive impact of presenting datasets in an interactive and visually compelling way, which enhances awareness and accessibility to those kinds of public data that

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3405656

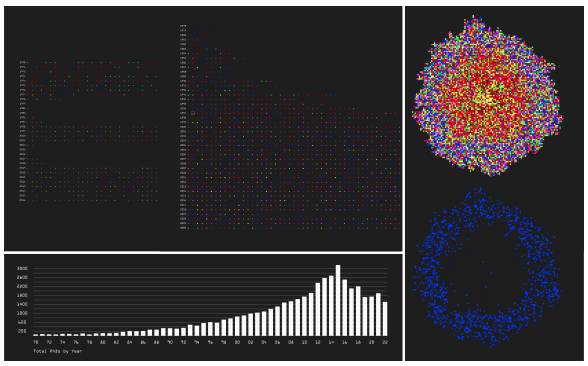

FIGURE 9. Snapshots of some insights highlighted with the visualization. Production gaps between years (top-left), production decrease from 2015 onwards (top-left), and emerging field over time (right).

would otherwise remain oblivious to the public.

P.ERMANENT H.ERITAGE D.ISPLAY

Data physicalization allows us to educate a broader audience by improving the efficiency of information retrieval and the memorability of data compared to similar designs shown only in flat, passive displays [23]. Ultimately, they can increase the user time spent on the site or the interest in interacting with the data visualizations due to the appeal of data visualizations themselves and their transformative potential of the user experience [24]. Moere [25] further emphasizes the significance of situating visualizations within their intended context, especially in public spaces. In that sense, a permanent installation is under development to be placed at one UNESCO World Heritage region of the city of Coimbra, at the University of Coimbra, the oldest university of Portugal (established on 1 March 1290) and one of the oldest of Europe. More specifically, it will be established next to the Tower of the University of Coimbra, at the Pátio and Paço das Escolas, formed called of Royal Palace of Alcáçova, which has housed the University since 1537 (Fig. 10).

Although the visualization was well-received, we recognize that the feedback we received came from a limited audience, which was not enough to fully validate the success of the work or its assumptions when exposed to the target audience. Placing this installation in such a location will allow a valuable opportunity to conduct a thorough on-site assessment by gathering more diverse and critical feedback from a broader audience. This evaluation could offer rich insights into the long-term impact of the installation, and its effectiveness in engaging and educating the public.

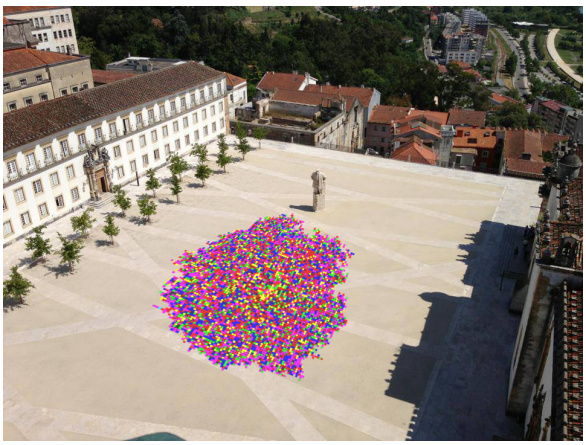

FIGURE 10. Model of the permanent installation to be installed at the Pátio and Paço das Escolas of the University of Coimbra using the algorithm of our work to represent all the doctorates concluded at the University of Coimbra (1970 -2022).

Additionally, it provides an opportunity to explore how public installations can effectively address specific topics when strategically positioned in locations relevant to the displayed data, as underscored by Moere [25].

The installation will represent all the doctorates carried out at the university to date and will be in constant evolution, being updated with the addition of a new element whenever a new doctorate is concluded. The installation elements will be produced with wooden materials. The foundation of this installation is to perpetuate the purpose of this work beyond digital media, promoting tangible dissemination of this work subject, nonetheless, in a historic area of the city with a large influx of academic and tourist communities which can only help spread such information throughout the country and even beyond borders.

CONCLUSION

The interactive installation "PhDs – Portugal has Doctors" represents a concerted effort to disseminate the extensive academic contributions originating from Portugal, namely through doctoral theses. The installation visually portrays all doctoral theses completed and recognized in Portugal from 1970 to 2022 in a manner that is both simplistic and abstract but that provides, at the same time, the opportunity to conduct analytical explorations to retrieve more detailed information from the data. These exploration of this data provided a comprehensive overview of the current state of doctoral education, offering insights into the

areas of growth and the demographic characteristics of PhD holders. Its reception has been notably positive, eliciting considerable engagement and prompting sensations of wonder and surprise, precisely as intended. Moreover, initial feedback suggests that the depiction of this academic landscape has unveiled previously unrecognized realities, even among members of the scientific community. All these observations indicate that there is an increasing potential for exploring new paths of data dissemination and analysis through alternative visualization and interaction approaches.

In future work, we want to conduct a deeper exploration of the dataset at hand. We pretend to improve the dataset with information about doctoral advisors and incorporate keywords associated with each thesis. While this endeavor may be time-consuming, it holds promise for uncovering valuable relationships between advisors and students, as well as intriguing connections between different scientific domains. Another possible exploration involves crossing the dataset with other additional datasets, such as socioeconomic or employment datasets, which could help to understand the current employment situation of the presented doctorates and the relationship between their academic pursuits and employment conditions at the national level.

Furthermore, we intend to delve into more granular statistical analyses. This includes identifying the most frequently represented foreign universities and countries among doctoral candidates who completed their studies abroad, allowing for comparisons with doctoral statistics from other nations. Moreover, we seek to explore correlations with social and economic events to better understand how these factors have influenced doctoral numbers over time. For instance, we are intrigued by the potential influence of historical events such as the establishment of universities or the societal shifts following the 1974 Carnation Revolution. Through these efforts, we seek to deepen our understanding of Portugal's academic landscape and shed light on its broader societal implications.

ACKNOWLEDGMENTS

This work is funded by national funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project CISUC - UID/CEC/00326/2020 and by European Social Fund, through the Regional Operational Program Centro 2020. The first author is funded by Fundação para a Ciência e Tecnologia under the grant SFRH/BD/144283/2019.

REFERENCES

1. P. Videira, "O mercado de trabalho dos doutorados em portugal e seus desafios," *Revista Portuguesa de Educaç ao*, vol. 35, no. 1, p. in press, 2022.
2. R. M. Gomes, L. Cerdeira, J. T. Lopes, H. Vaz, R. Brites, B. Cabrito, S. Silva, T. Patrocínio, M. d. L. Machado-Taylor, P. Peixoto *et al.*, "Working paper 4: focus groups: general conclusions," Centro de Investigação e Estudos de Sociologia (CIES-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Tech. Rep., 2014.
3. P. Teixeira and P. Videira, "A tale of expansion and change: major trends in doctoral training and in the doctoral population in portugal," *Doctoral Education for the Knowledge Society: Convergence or Divergence in National Approaches?*, pp. 67–87, 2018.
4. N. Costa, R. Costa, A. Borges, V. Branco, R. Cunca, A. Modesto, and A. C. Silva, "Mapping the research thread of phds in design: a phd citation analysis of the portuguese doctorates," in *Design for Tomorrow—Volume 2: Proceedings of ICoRD 2021*. Springer, 2021, pp. 203–216.
5. J. Walny, S. Storteboom, R. Pusch, S. M. Hwang, S. Knudsen, S. Carpendale, and W. Willett, "Pixelclipper: Supporting public engagement and conversation about visualizations," *IEEE Computer Graphics and Applications*, vol. 40, no. 2, pp. 57–70, 2020.
6. E. M. Peck, S. E. Ayuso, and O. El-Etr, "Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 2019, Glasgow, Scotland, UK, May 04-09, 2019*, S. A. Brewster, G. Fitzpatrick, A. L. Cox, and V. Kostakos, Eds. ACM, 2019, p. 244. [Online]. Available: https://doi.org/10.1145/3290605.3300474
7. J. Boy, A. V. Pandey, J. Emerson, M. Satterthwaite, O. Nov, and E. Bertini, "Showing people behind data: Does anthropomorphizing visualizations elicit more empathy for human rights data?" in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, May 06-11, 2017*, G. Mark, S. R. Fussell, C. Lampe, m. c. schraefel, J. P. Hourcade, C. Appert, and D. Wigdor, Eds. ACM, 2017, pp. 5462–5474. [Online]. Available: https://doi.org/10.1145/3025453.3025512
8. L. A. de Macêdo Morais, Y. Jansen, N. Andrade, and P. Dragicevic, "Showing data about people: A design space of anthropographics," *IEEE Trans. Vis. Comput. Graph.*, vol. 28, no. 3, pp. 1661–1679, 2022. [Online]. Available: https://doi.org/10.1109/TVCG.2020.3023013

9. F. Samsel, “Art-science-visualization collaborations; examining the spectrum,” *Proceedings of the IEEE VIS arts program (VISAP)*, 2013.
10. N. Miebach, B. D. Campbell, and F. Samsel, “Nathalie miebach: Sculpted data infused with craftsmanship,” *IEEE Computer Graphics and Applications*, vol. 42, no. 1, pp. 7–16, 2022.
11. P. Cruz, “Wrongfully right: applications of semantic figurative metaphors in information visualization,” *IEEE VIS Arts Program (VISAP)*, pp. 14–21, 2015.
12. F. B. Viégas and M. Wattenberg, “Artistic data visualization: Beyond visual analytics,” in *International Conference on Online Communities and Social Computing*. Springer, 2007, pp. 182–191.
13. R. van Koningsbruggen and E. Hornecker, ““it’s just a graph" - the effect of post-hoc rationalisation on infovis evaluation,” in *Proceedings of the 13th ACM SIGCHI Conference on Creativity and Cognition, C&C 2021, Virtual Event / Venice, Italy, June 22-23, 2021*, C. Sas, N. A. M. Maiden, B. P. Bailey, C. Latulipe, and E. Y. Do, Eds. ACM, 2021, pp. 33:1–33:10. [Online]. Available: https://doi.org/10.1145/3450741.3465257
14. B. Lee, E. K. Choe, P. Isenberg, K. Marriott, and J. Stasko, “Reaching broader audiences with data visualization,” *IEEE Computer Graphics and Applications*, vol. 40, no. 2, pp. 82–90, 2020.
15. T. Nagel, C. Pietsch, and M. Dork, “Staged analysis: From evocative to comparative visualizations of urban mobility,” in *2017 IEEE VIS Arts Program (VISAP)*, 2017, pp. 1–8.
16. L. Manovich, “Cultural analytics: Visualizing cultural patterns in the era of “more media”,” *Domus March*, 2009.
17. N. Cawthon and A. Vande Moere, “Qualities of perceived aesthetic in data visualization,” in *Conference on Designing for User Experiences*. ACM, 2007.
18. A. Ramirez Gaviria, “When is information visualization art? determining the critical criteria,” *Leonardo*, vol. 41, no. 5, pp. 479–482, 2008.
19. D. Baur, M. Sedlmair, R. Wimmer, Y. Chen, S. Streng, S. Boring, A. Butz, S. K. Feiner, A. M. MacEachren, and S. Jánicke, “Trends in information visualization,” *arXiv preprint arXiv:1802.06864*, 2018.
20. Q. Li, “Data visualization as creative art practice,” *Visual Communication*, vol. 17, no. 3, pp. 299–312, 2018.
21. A. Bundy and L. Wallen, “Breadth-first search,” *Catalogue of artificial intelligence tools*, pp. 13–13, 1984.
22. S. J. Stroessner, J. Benitez, M. A. Perez, A. B. Wyman, C. M. Carpinella, and K. L. Johnson, “What’s in a shape? evidence of gender category associations with basic forms,” *Journal of Experimental Social Psychology*, vol. 87, p. 103915, 2020.
23. Y. Jansen, P. Dragicevic, and J.-D. Fekete, “Evaluating the efficiency of physical visualizations,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’13. New York, NY, USA: Association for Computing Machinery, 2013, p. 2593–2602. [Online]. Available: https://doi.org/10.1145/2470654.2481359
24. L. Abbate and C. Marino, “Designing data interaction in exhibitions contexts,” in *Proceedings of DRS*. Design Research Society, jun 2022. [Online]. Available: https://doi.org/10.21606/drs.2022.693
25. A. V. Moere and D. Hill, “Designing for the situated and public visualization of urban data,” *Journal of Urban Technology*, vol. 19, no. 2, pp. 25–46, 2012. [Online]. Available: https://doi.org/10.1080/10630732.2012.698065

Pedro Silva is a cross-media designer, researcher and PhD candidate at the Department of Informatics Engineering of the University of Coimbra with publications and installations across multidisciplinary conferences such as IV, IVAPP, xCoAx, and SIGGRAPH. Contact him at pedros@dei.uc.pt.

Pedro Martins is a researcher at CISUC and an Associate Professor at the Department of Informatics Engineering of the University of Coimbra. He holds a PhD degree in Computer Science from the University of Coimbra and his main research interests are in the fields of Computational Creativity, Computer Vision, Computer Graphics, and Computational Art. Contact him at pjmm@dei.uc.pt.

Penousal Machado is an Associate Professor at the Department of Informatics of the University of Coimbra, the coordinator of CISUC’s Cognitive and Media Systems group, and the president of SPECIES society. His work was featured in the Leonardo journal, Wired magazine and presented in venues such as the National Museum of Contemporary Art (Portugal) and the “Talk to me” exhibition of the Museum of Modern Art, NY (MoMA). Contact him at machado@dei.uc.pt.