xai-primer.com — A Visual Ideation Space of Interactive Explainers

Beatrice Gobbo
Politecnico di Milano
Milano, Italy
beatrice.gobbo@polimi.it

Tommaso Elli
Politecnico di Milano
Milano, Italy
tommaso.elli@polimi.it

Uta Hinrichs
University of Edinburgh
Edinburgh, UK
uhinrich@ed.ac.uk

Mennatallah El-Assady
ETH AI Center
Zurich, Switzerland
melassady@inf.ethz.ch

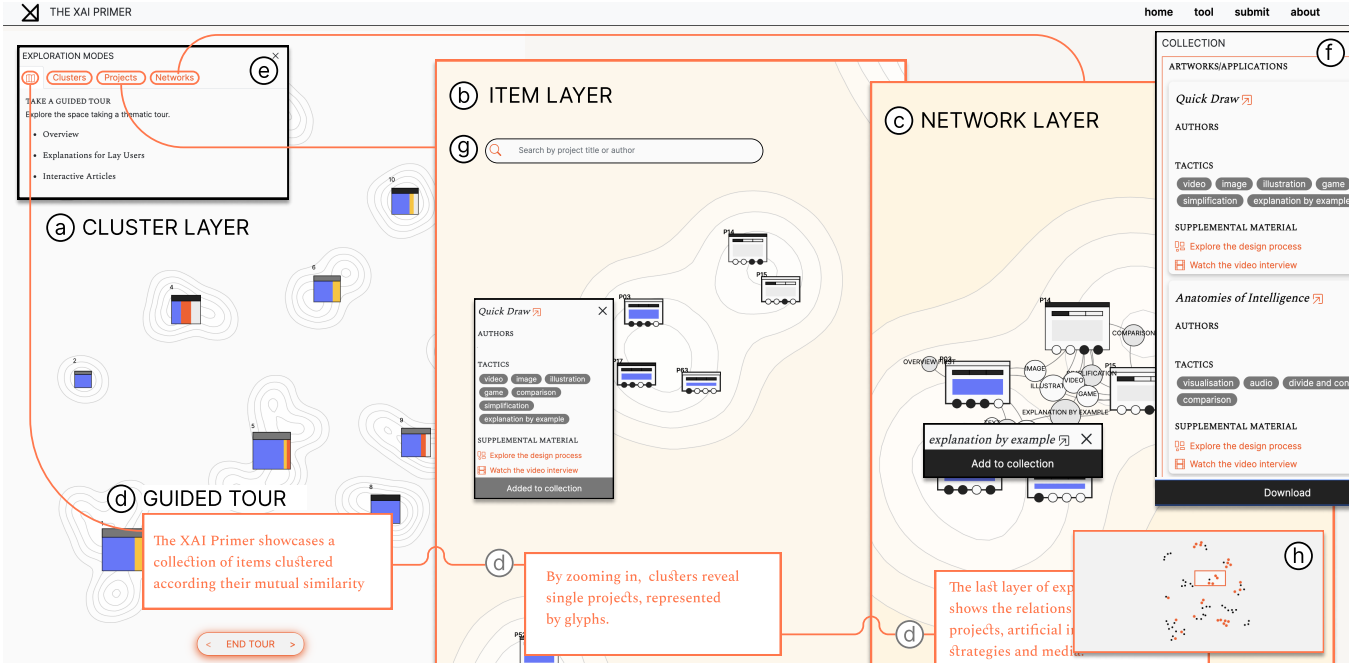

Figure 1: A summary of the main functions of the XAI Primer. The design space is structured into three layers, namely: the cluster layer (a), the item layer (b) and the network layer (c) that can be explored in an open-ended and serendipitous way. Guided tours (d) provide focused reading of the space by automatically panning and zooming. The exploration modes panel (e), the collection panel (f), the search bar (g) and the minimap (h) complete the interface, respectively allowing users to navigate across views, collect and download lists of items, search for authors or projects, orient them in space and highlighting the same items in the space.

KEYWORDS

Explainable Artificial Intelligence, Open Ended Exploration

CHI '22 Extended Abstracts, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9156-6/22/04.
https://doi.org/10.1145/3491101.3519880

ACM Reference Format:
Beatrice Gobbo, Tommaso Elli, Uta Hinrichs, and Mennatallah El-Assady. 2022. xai-primer.com — A Visual Ideation Space of Interactive Explainers. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3491101.3519880

1 INTRODUCTION

A recent branch of Explainable Artificial Intelligence (XAI) is concerned with the design of visual interactive explainers to make AI systems understandable. However, because of the field's quick evolution and interdisciplinary nature, it is challenging to have an

overview and develop new approaches. XAI researchers have made several attempts to organize the explaining process into conceptual frameworks. Barredo Arrieta et al. focused on XAI concepts and taxonomies for Deep Learning [1], while others organized surveys according to target audiences, visual models and tasks. [11, 15]. The recent EUCA project by Jin et al. made a preliminary effort in defining strategies for addressing explanations to non-experts [18]. Our work is based on contributions made at the intersection of visualization and the humanities/cultural heritage, which have seen advancements in the design of interactive tools aimed at overcoming traditional interface design approaches to cultural collections, such as grid-based interfaces centered around targeted search. [10], supporting visual and open-ended exploration [22]. In this context, *serendipity* [5, 20] and *generosity* [21] have been considered as guidelines to design interfaces for browsing cultural collections.

We present the XAI Primer, a digital interface that organizes XAI research and projects by *explanation strategy*, *media*, and *usage scenarios* at different levels of granularity and allows users to browse the space by blending open-ended exploration with guided tours using a museum metaphor to arrange elements in space and inform the design process of interactions.

The XAI Primer is a visual ideation space where designers can find inspiration and case studies even outside their academic field. Indeed, the XAI Primer juxtaposes XAI elements from different communities, positioning them according to their similarity in terms of strategies and used techniques. Hence, it provides a comprehensive and interdisciplinary overview of the state-of-the-art. Therefore, the XAI Primer can be considered a growing design space that reflects on past, current and future efforts in XAI. Since the interface has no limited boundaries, it is our intention to promote the XAI Primer as an ever-growing tool, subject to change and able to accommodate new **items** suggested by users.

The CHI community is interdisciplinary in nature and brings together experts in different fields who could benefit from the exploration of the XAI Primer and contribute with new proposals. Indeed, CHI community has been already committed for years now to encouraging "holistic" perspectives in XAI [7] and innovative solutions to support AI explanations [4].

2 THE XAI PRIMER DESIGN SPACE

First and foremost, the XAI Primer supports researchers working on visual interactive explainers to ideate unexplored combinations of strategies and techniques. The interface supports three main tasks, namely (1) **overview** of the space; (2) **exploration** and comparison of XAI items; and (3) **analysis** of individual items and their design context, such as raw materials, doodles and interviews with authors.

The XAI Primer is tailored to support three stakeholder groups, namely (1) **authors** (of featured items) in framing, summarizing and categorizing existing XAI items; (2) **curators** (e.g., conference organizers and program managers) in collecting items and developing guided tours through the space; and (3) **designers** (hopefully prospective authors) in exploring XAI items for inspiration and ideation.

We came up with three main design goals, aligned to the three stakeholder groups. The XAI Primer allows *authors* (1) **to summarize and categorize items**, supporting them in identifying their work in the context of similar related works. Moreover, it enables curators, namely experts researchers in the field dittoing the role of the curator within museum exhibitions, to (2) **inform new collections of items** making connections between existing works, and identifying new trends. Finally, as its first aim, supports designers in (3) **ideating new items**.

3 COLLECTING ITEMS' METADATA

The XAI Primer currently includes a collection of 71 Explainable Artificial Intelligence projects gathered from the VISxAI Workshop [9], Distill [3], Google Arts Experiments [12], IEEVIS Art Program [13], as well as independent artworks from the fields of art, media studies, communication design and data activism [14].

In order to generate a set of metadata that could be applied across items to highlight their similarities and differences and facilitate the visual representation of this research space and its interactive exploration, we coded each item according to the XAI-Building Blocks (XAI-BB) Framework [8]. This framework presents different dimensions that have been derived based on explanation strategies discussed in the fields of pedagogy, philosophy, and communication studies in order to characterize XAI processes.

4 WHAT ATTENDEES WILL EXPERIENCE

The XAI Primer [1] in its current form is a web-based platform, primarily consisting of an interactive prototype [2] where XAI projects, strategies and artworks are presented (Figure 1). Besides, an About page contains information about our design process and a Submit page allows visitors to propose new case studies or strategies. After many design iterations and interactive mock-ups, the XAI Primer has been implemented using a combination of D3.js [2] and React.js [16]. As previously stated, the XAI Primer uses analogies of a landscape and a museum setting (more spatially constrained) to graphically display the interactive explainers. The location of items is determined by their similarity and is accomplished by applying the Uniform Manifold Approximation that compresses the multi-dimensional data space generated by our qualitative coding into a 2D vector space (UMAP) [19].

4.1 Three layers of exploration

Three levels of granularity—**clusters**, **items**, and **networks** structure the interface. The background is populated with textual annotations and non-graphical signs that serve viewers as a preliminary orientation tool. Glyphs encode information and fluidly adapt to the level of granularity the visitor is exploring, populate the interface. Indeed, encoding many dimension in a single glyph is challenging, so only selected dimensions are shown at the cluster, item and network level, progressively presenting a more and more complete picture of an XAI item. (See Figure2)

Cluster Layer. The cluster layer is the first layer visible the visitor can interact with. It provides an overview of XAI items, positioning and grouping them according to their characteristics using UMAP. To follow the museum metaphor, clusters of items can be considered as rooms in a museum where similar pieces are

[1] https://xai-primer.com/
[2] https://xai-primer.com/tool

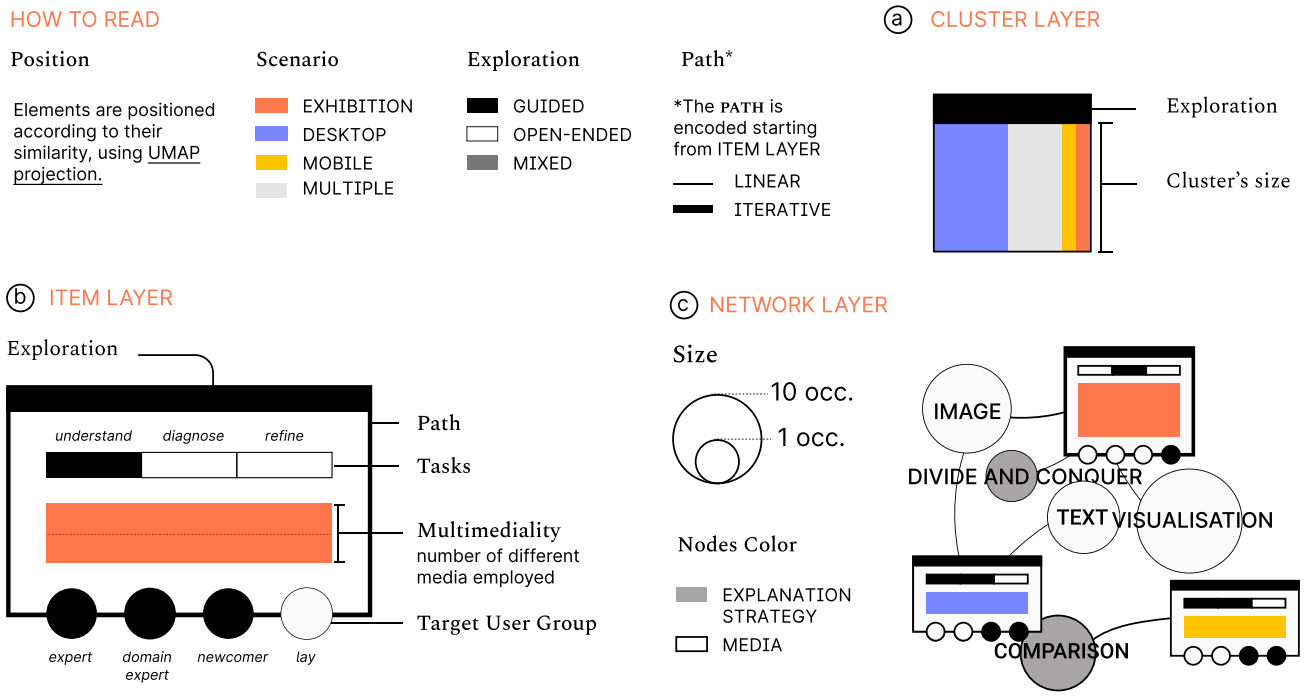

Figure 2: An overview of the glyphs at different levels of granularity. The cluster layer (a) presents treemap-like glyphs, encoding desktop and exhibition as dimensions. In the item layer the task, media, target user group and path dimensions are added to glyphs. Finally, in the network layer, relationships between XAI items within the same cluster are revealed.

showcased for visitors to further explore. At this level, XAI items are grouped and represented as treemap-like glyphs, where size represents the number of items featured in each cluster, and color represents the *scenario of use*. The top bar represents the type of expected *exploration*.

Item Layer. Zooming into the XAI Primer reveals the item layer, which provides information about individual items. At this level, glyphs no longer represent clusters of XAI items, but individual items. Elements of the cluster view glyph remain, but others are introduced to unfold the item details and to facilitate the identification of similarities and differences between individual items.

Network Layer. Finally, zooming in further, the network layer reveals the hidden relationships among items, overlaying across the cluster a graph connecting media and explanation strategies. The size of each node represents the number of occurrences, thus the larger the node, the more that explanation strategy (dark-grey nodes) or media (white nodes) is used within the cluster being observed. We use a force-directed layout to position keyword nodes [17]. Combining these different layers through common interaction techniques such as panning and zooming, the XAI Primer allows to fluidly combine vertical and horizontal exploration of XAI items [6], where visitors can first get an overview of item clusters before diving into a more in-depth exploration of individual contributions to XAI.

Item Details. Visitors can explore individual details about an XAI item, by selecting its representing glyph directly from the Item or Network layer. Selecting an item will bring up a dedicated page presenting details about the item, including its authors, a brief description of its focus, as well as Building Blocks and related attributes. Where possible, we also include information about the design process that authors went through when designing the item.

4.2 Modes of Exploration Supported in the XAI Primer

The XAI Primer supports different types of exploration— **open-ended** exploration, **guided tours** that allow the traversal of the design space alongside curated perspectives, and **targeted search** of individual items of interest.

Open-Ended Exploration. Visitors of the XAI Primer can fluidly explore the design space by zooming in and out of the layers. Through the open-ended exploration of the space, we aim to support serendipitous discoveries, drawing from design considerations discussed by Thudt et al. [20]. Therefore, layered visual overviews provides a generous [21] entry point into the collection where visitors can start their explorations without knowledge of the design space or knowing what to look for. Visitors can start exploring guided by their curiosity and interests, without following any prescriptive path. For example, visitors can start their exploration by browsing the space at the *cluster level*, reading annotations and digging in specific areas. However, they can also use the exploration modes panel to directly jump to the network view to explore individual items and their thematic links. These modes of exploration

make clear our intention to provide visitors with a space that is as usable as a museum exhibit, where one can consult a floormap firstly or directly explore thematic rooms. On the bottom right of the interface, a *minimap* facilitate open-ended explorations by presenting a small replica of the item layer which acts as a wayfinder, preventing visitors from getting lost. Moreover, when exploring the network layer clicking on a keyword node, other portions of the design space are highlighted where related items can be found. In this way, new connections between XAI items across the design space are made visible, again allowing for the discovery of new and unexpected links.

Guided tours complement open-ended explorations, that allows visitors to browse the space following a curated path. At the current stage, the XAI Primer includes three guided tours. The first provides an overview or the space, while the others guide visitors to dominant trends or anomalies within the collection.

Targeted Search. Of course, there are cases where a visitor may want to find individual XAI items directly. This type of targeted search is supported through a common *search bar,* where visitors can search the design space by author or item title.

Collecting, Connecting & Inspecting. In order to enable visitors to keep track of XAI items of interest they may have found as part of their explorations, the XAI Primer features the *collection panel*. The collection panel reminds of a citation or a reference manager tool where information about XAI items can be downloaded in .txt format to inform both the practical design of XAI projects, as well as the framing of XAI projects in the context of related work as part of publications.

5 DISCUSSION

A qualitative study conducted on a sample of 9 experts representing XAI Primer end-users provides a first glimpse into the potential of the XAI Primer as a tool for *reflection* on topics and trends in the field of XAI, as well as how work in this relatively young community is shaped by different disciplines and approaches spanning the sciences and arts. Our findings are particularly encouraging in terms of the Primer's potential as a *generative* tool that informs not only the design of new XAI approaches, but also facilitates the framing of XAI research in the light of previous work. Finally, our findings indicate the potential, in particular, of the guided tours to invite and help acclimatize new audiences to the field of XAI, enabling them to *learn* about existing approaches and their relations. The XAI Primer supports researchers and practitioners in the *ideation* phase. The distribution of items' details across layers makes the content digestible, clear, and helpful in seeking inspiration, comparing approaches, probing the state-of-the-art, and collecting cases. Horizontal and vertical exploration, facilitated by fluid interaction and guided tours, make the space open and elastic, ready to accommodate brand-new items, interactions, and stories to be narrated.

REFERENCES

[1] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. Explainable Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58 (2020), 82–115. https://doi.org/10.1016/j.inffus.2019.12.012 arXiv:1910.10045

[2] Mike Bostock. 2021. Data-Driven Documents. https://d3js.org/ Accessed: 2021-09-09.

[3] Shan Carter and Chris Olah. 2021. Distill. https://distill.pub/journal/ Accessed: 2021-09-09.

[4] Chi201. 2019. Human-Centered Machine Learning Perspectives Workshop. https://gonzoramos.github.io/hcmlperspectives/index.html Accessed: 2022-01-01.

[5] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. 2011. The information flaneur: a fresh look at information seeking. In *Proceedings of the International Conference on Human Factors in Computing Systems, CHI 2011, Vancouver, BC, Canada, May 7-12, 2011*, Desney S. Tan, Saleema Amershi, Bo Begole, Wendy A. Kellogg, and Manas Tungare (Eds.). ACM, Vancouver, BC, Canada, 1215–1224. https://doi.org/10.1145/1978942.1979124

[6] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. 2012. Fluid Views: A Zoomable Search Environment. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Capri Island, Italy) *(AVI '12)*. Association for Computing Machinery, New York, NY, USA, 233–240. https://doi.org/10.1145/2254556.2254599

[7] Upol Ehsan, Philipp Wintersberger, Q. Vera Liao, Martina Mara, Marc Streit, Sandra Wachter, Andreas Riener, and Mark O. Riedl. 2021. Operationalizing Human-Centered Perspectives in Explainable AI. , 6 pages. https://doi.org/10.1145/3411763.3441342

[8] Mennatallah El-Assady, Wolfgang Jentner, Rebecca Kehlbeck, Udo Schlegel, Rita Sevastjanova, Fabian Sperrle, Thilo Spinner, and Daniel Keim. 2019. Towards XAI : Structuring the Processes of Explanations.

[9] Mennatallah El-Assady, Duen Horng (Polo) Chau, Adam Perer, Hendrik Strobelt, and Viegas Fernanda. 2021. Workshop on Visualization for AI Explainability. http://visxai.io/ Accessed: 2021-09-09.

[10] EU. Europeana Collections. Europeana Collections. https://classic.europeana.eu/portal/en

[11] Rafael Garcia, A. Telea, B. Silva, J. Tørresen, and J. Comba. 2018. A task-and-technique centered survey on visual analytics for deep learning model engineering. *Comput. Graph.* 77 (2018), 30–49.

[12] Google. 2021. Google Arts & Culture Experiments. https://artsexperiments.withgoogle.com Accessed: 2021-09-09.

[13] Yoon Chung Han and Charles Perin. 2021. The IEEE VIS 2021 Arts Program. http://visap.net/ Accessed: 2021-09-09.

[14] Fred Hohman, Matthew Conlen, Jeffrey Heer, and Duen Horng Chau. 2020. Communicating with Interactive Articles. https://doi.org/10.23915/distill.00028

[15] Fred Hohman, Minsuk Kahng, Robert Pienta, and Duen Horng Chau. 2019. Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers. *IEEE Transactions on Visualization and Computer Graphics* 25 (2019), 2674–2693. https://doi.org/10.1109/TVCG.2018.2843369

[16] Facebook Inc. 2021. React. A JavaScript library for building user interfaces. https://reactjs.org/ Accessed: 2021-09-09.

[17] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE* 9, 6 (06 2014), 1–12. https://doi.org/10.1371/journal.pone.0098679

[18] Weina Jin, Jianyu Fan, D. Gromala, P. Pasquier, and G. Hamarneh. 2021. EUCA: A Practical Prototyping Framework towards End-User-Centered Explainable Artificial Intelligence.

[19] L. McInnes, J. Healy, and J. Melville. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 [stat.ML]

[20] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. 2012. *The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization.* Association for Computing Machinery, New York, NY, USA, 1461–1470. https://doi.org/10.1145/2207676.2208607

[21] Mitchell Whitelaw. 2015. Generous Interfaces for Digital Cultural Collections. *Digital Humanities Quarterly* 9, 1 (2015), 1–16.

[22] Florian Windhager, Paolo Federico, Gunther Schreder, Katrin Glinka, Marian Dörk, Silvia Miksch, and Eva Mayr. 2019. Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges. *IEEE Transactions on Visualization and Computer Graphics* 25, 6 (2019), 2311–2330. https://doi.org/10.1109/TVCG.2018.2830759