

Some of this is excellent, some is much less so. But over all, good job -

Data physicalization

Group 6

Kevin HUART
Mathilde VERRIEZ CHASTANG
kevin.huart@universite-paris-saclay.fr
mathilde.verriez-chastang@universite-paris-saclay.fr
Université Paris-Saclay
Orsay, France

Idriss TAGHOURY
Antoine AYOUB
idriss.taghoury@universite-paris-saclay.fr
antoine.ayoub@universite-paris-saclay.fr
Université Paris-Saclay
Orsay, France

ABSTRACT

This review ~~aims to resume the work of many researcher teams~~ *focuses on the advantages of* around data physicalization. ~~We are passing through the advantages of physicalization, which are often emphasized by researchers.~~ *including* data physicalization enhances data perception, cognitive abilities, learning, memorization. ~~Data physicalizations also provide a better understanding of the data they represent. Studies show that non-technical users who might lack the knowledge to interpret data find the simplification of said data through its physicalization impactful on the way they perceive and understand it. But it comes with a major drawback: physical representations of data can be difficult to obtain: design issues, cost, materials, lack of preparation...~~ *improves how* Researchers are aware of this and try to provide guidelines on how to systematically create relevant data physical models according to the context they are placed in and the data they are referring to. There is also a possibility for such models to autonomously update to data and context: dynamic data physicalizations. Dynamic data physicalization is more and more studied by some researchers who think it provides new ways of collaborating and making decisions. Finally, we discuss the future directions this area of research could take in the years to come.

brought by data physicalization and proposed designs, solutions and discussions about this subject. ~~In this paper, we communicate a small overview of the current state of the art in the area of data physicalization and more specific subtopics.~~ *summarizes a subset of*

2 RELATED WORK

Data physicalization is mostly seen by researchers as a promising field with lots of opportunities to exploit [7]. In this section we talk about the domains of data physicalization that seem to attract researchers interest and their contributions to these domains. First, researchers ultimately try to enhance the ways of customizing, interpreting and interacting with data through data physicalization. But it eventually brings about conception and implementation challenges. To counter that, researchers tend to look *for a baseline* on (1) what Data Physicalization is and (2) how to systematically design them according to the context in which they are used and to how users perceive the corresponding data. Finally, some researchers are interested in exploring how data physical representations can be made autonomous and dynamic and how it could benefit users [11, 18].

informed style
You can shorten content is good

1 INTRODUCTION

Data Physicalization is *a field of research* that uses physical objects to represent, visualize and communicate data. Some researchers justify the emergence of this field with the fact that mankind always used tangible objects as physical information representations [7, 11]. These researchers, ~~as well as many others who devoted their work to this matter, identified several problems [3, 4, 17] that led them to contribute to this field. They deplore the fact that current techniques of data visualization need to be revised; they either require too much technical knowledge for the data to be interpreted correctly [12] or do not provide efficient ways of enhancing collaboration [11]. Even though progress has been made in the field and research has been conducted, researchers still point out that conception and design guidelines for data physicalization are missing, making it difficult for different physical data representations to adapt from one data set to another [16], or for researchers new to the field of data physicalization to present efficient solutions [4]. Other researchers started insisting on the limitations of some types of data physicalizations, opposing passive and dynamic data physicalizations, stating that passive data physicalizations could hardly autonomously update when the visualized data sets require it [7, 11, 18]. As one can see, researchers identified the challenges~~ *Nice intro*

2.1 Enhancement of data interpretation and customization

Some researchers think the physical representation of data can have a major advantage compared to their virtual counterparts. This section illustrates how researchers see in Data Physicalization a way of improving interaction with data.

Jansen et al. classify the benefits of Data Physicalization in three categories : perceptual, cognitive and societal [7]. First, Physicalization better exploits the user's active perception skills by physically manipulating the representation, which is not possible with a visual representation, and the spatial perception skills with a better depth perception on physical 3D data than virtual one. Physicalization can also be enhanced with non-visual senses, such as haptic [2, 9] or sound [15] information that can be used in an intermodal approach. Second, given the benefits of physical manipulation in learning and education, they expect Data Physicalization to also improve cognition and learning. Raidou et al. [13] conducted studies which results show that manipulating physical models representing data proves to be more engaging than with only virtual visualization and that some complex notions could be even easier to understand through physical manipulation. *Like Simons work* Likewise the work of Ang et al. [1] on the visualization of blood flow data *emphasize the fact that* engaging one's sense of touch with physical models when manipulating data

However,

point to the lack of

Reword

... Avoid ellipses in technical papers

makes depicting complex patterns easier. Last, Data Physicalization also makes data more accessible for a broader audience, including visually-impaired people and is believed to better engage people with the data. Additionally, physicalization can also be used for a personal [9, 19] or artistic [14, 20] purpose.

Willett et al. [20] take another approach and propose to compare different techniques of data representation. They present situated visualizations, situated physicalizations, embedded visualizations and embedded physicalizations as four ways of representing data. They compare these data representation techniques by designing a scenario during which these techniques would each be used to perform the same task. They also try to describe the relationship between the data and the physical token it is referring to. They eventually arrive at the conclusion that embedded physicalizations allow for more possibilities in how analysts interact with data, but point out a major drawback being the implementation difficulty of such a technique. Finally, they propose an overview of how modern technology could apply to satisfy different data representation techniques. Although, they provide some insight about different data representation techniques, they did not conduct any experiment to solidify their comparison.

By contrast, Jansen et al. propose two experiments in order to compare virtual and physical representations [6]. The principle being to study the speed and the error rate of several tasks performed on a bar graph. These two experiments reveal that the physical representation is much more efficient since the working time on a physical representation is at least 20% less than on a screen representation. In addition, being able to touch the sculpture offers a higher capacity for understanding, but they also state that this solution of illustration is quite hard to realize without good methods and tools.

Data physicalization can be useful in everyday life. Thudt et al. [19] show a study of the advantages offered by the creation of visualizations on "personal" problems. The study begins with a design phase in which the participants have to choose a theme and create a corresponding data summary (decisions on data attributes, possible categorizations, scales and possible range of values of attributes). The study was carried out over the course of four weeks, the majority of the participants claimed that reading the data was clear and fun and it allowed them to be more critical of their action and even find some form of motivation. The act of creating their own data physicalization also allowed the participant to become more aware of their problems. By the end of the study, the majority of participants had solved their problem. Data physicalization can therefore be a good asset on a daily basis by transforming one's actions into a data set that can be analyzed and interpreted simply.

The act of being in control of the creation process allows to model the data with materials and colors appreciated by the user. Much like Thudt et al., Perovich et al. [12] emphasize the capacity of data physicalization to be appealing to non-technical users, who might have a hard time comprehending data sets. In their article, they claim that despite government institutions being more and more transparent on the data they share, said data can be hard to interpret for common users. As mentioned, Perovich et al. explore the solutions offered by data physicalization to tackle this issue. Their implementation of *Chemicals in the Creek* stimulated enthusiastic reactions from the targeted non-technical users who were

mostly able to understand the data.

Avoid this phrase

There are also some pitfalls because of which some inexperienced physicalists could get stuck in the creation process of physicalization [4]. In this research, Huron et al. conduct an experiment during workshops at *Futur en Seine 2014* and *Twente 2016*. During these, they have noticed three different pitfalls. The first one occurred during the data preparation stage, the second at the ideation step and the third at the building stage. For the first two, participants took way too much time to respectively prepare data - may be due to a lack of preprocessing - and to imagine how to represent it. For the last, a too ambitious building project, with largely too complex construction, seems to be the reason why participants get stuck. This article as well as some others [7, 16] deplore the fact that the field of data physicalization lacks design guidelines to help researchers create efficient data physicalizations.

2.2 Conception and design

This section presents an overview of the different approaches around data physicalization conception.

Jansen et al. [7] highlight three challenges in designing perceptually effective data physicalization : (1) understand the design space of the representation, (2) understand how different approaches affect the perceptual effectiveness of the representation of data and (3) find how to implement the representation. In order to understand the design space, they advise to first understand the physical variables that want to be represented (such as smoothness, hardness or sponginess) by combining the different sensory variables (visual, haptic, etc.) usually associated with the physical variable in the real world.

Stusak et al. [16] provide further information about the influence of 14 variables on what 3D data physicalizations look like and how they are perceived. They group these variables into four categories: geometric variables (position, orientation, global shape, exact shape), color variables (hue, saturation, luminance, optics), tactile variables (roughness, lay, temperature, compliance), and kinesthetic variables (slipperiness, weight). They discuss the performance of these variables according to four characteristics (selective, associative, quantitative and order). Results show that most of the variables are selective, associative and provide a way of ordering data physicalization units, meaning that a change in any of these variables allows the user to differentiate one data representation from an other or regroup data representations easily. Although this paper provides relevant insight for the conception of data physicalizations, the authors only consider variables that have an impact on visual and haptic senses excluding sound and taste. They also only discuss the individual performance of the variables without mentioning the influence of mixing variables together on the perception of data physicalizations.

As for perceptual effectiveness, Jansen et al. [7] note that the two important factors are how to encode the data and what physical size to use in representing it. Although guidelines for encoding physical data are yet to be developed, such guidelines already exist for visualization and can be used as a starting point in developing guidelines for Physicalization. Finally, the implementation of

no capitals

Para graph

Much too much detail for a literature review - summarize the key points and tell us why they matter.

Avoid explore

This section needs work, on the writing

the physicalization needs to take into account the cost, replicability, fidelity, depletion and environmental impact for choosing appropriate materials to represent the physical variable.

Huron et al. also explore some limitations in the physicalization creation process [3]. They found that the definition of data units is a critical point. Indeed, choosing a unit or modifying data amplitude may be a major issue. The organization or rearrangement of data units in purpose of presenting a new data set is as well an important choice because of the physicalization readability. This has been confirmed by Jansen and Hornbæk [8] two years after, with their research on how the size of data physical artifacts can help expressing the data's value. They found that the "size" can be interpreted in different ways. They have based their research on two types of artifacts - spheres and bar charts - and defined that the most expressive "size" of a sphere is its surface while it is length for bar charts items.

Huron et al. provide further work on the subject [4] and share a method about the conception process of physicalization which is composed of five main steps: preparing the data (1), this means searching for convenient sets of data to be physicalized; then one has to imagine ways to assemble one's representation of data (2) and the materials one will have to use (3); after that, the main building task (4) can start and at the ending, one has to verify that what one just made reflects their thoughts (5). Previously, Huron et al. [3] were presenting a physicalization support tool and community named *InfoVis* which aims to democratize data physicalization around Computer Science communities. Meanwhile, Swaminathan et al. [17] introduced another tool named *MakerVis*. This one has the same objective as *InfoVis* but both are not making visualizations on the same support: *MakerVis* is about physical representation while the other concerns the virtual representations which can support physicalization.

Jansen et al. [5] refer to the notion of "PIPELINE INFOVIS" which is the sequence of data transformations changing raw data into physical representations. This sequence consists of three steps:

- Data Transformation: it is the processing of raw data to make it suitable for visualization. This can be obtained by filtering, concatenating data from several sources, making them compatible with the techniques used in the next steps
- Visual Mapping: it is the act of giving an initial visual form (generally abstract) to the data set. It corresponds in most cases to the construction of a graphic primitive
- Presentation Mapping: It is the act of transforming an abstract visual form into a terminal form (complete and comprehensible) which can be displayed, printed or manufactured. This step includes several possible operations such as the specification (specifies the last details of all the variables), the style (allows to unify the different parts of the whole visualization), the optimization (allows to facilitate the reading of a visualization) or decoration (makes it easier to read and interpret a visualization)

Data physicalization can also have an artistic purpose thanks to the Data Object [14]. A data object is an everyday object designed

thanks to data physicalization. Objects can range from household appliances and street furniture to educational materials and interactive exhibits. There are two methods to design such objects:

- Method 1
 - (1) Select the data sources, filter them and examine them
 - (2) Build initial 2D and 3D representations of the data
 - (3) Develop concepts for objects related to the data set
 - (4) Prototype and evaluate the power of the data object to transmit data when in use
- Method 2
 - (1) Select the type of object to design
 - (2) Develop data concepts related to the object
 - (3) Play with perceptual and conceptual mappings between data and object
 - (4) Prototype and evaluate the ability of the data object to transmit data when in use

Sosa et al. [14] present two designs to illustrate these methods: the first one consists of two tables. The volume of the table legs represents an economic disparity between Mexico and New Zealand. The table representing Mexico has very wide feet, which makes the user uncomfortable since he does not have enough room to be able to put his legs under the table. The principle is to encourage a conversation between the users of the tables so that they correlate their experience with the represented data in order to have a better understanding of it.

The second design is a table surrounded by chairs created to represent the inequality in the wealth distribution in the world. The data is: 20% of the population owns 82.7% of the world's wealth while the remaining 80% share only 17.3%, the poorest 40% owning virtually nothing. The table has been designed in such a way that there are 5 chairs, with one chair on one side and the others on the other side. The isolated chair owns 80% of the table while the four others have to share the rest (with two chairs having almost nothing).

Buur et al. also bring up the question of the physicalization of Big Data [2], as physical representation cannot capture the entirety of a large data set. Reducing data is a common practice in Data Physicalization in order to make the data more understandable, but the need for reducing in order to represent Big Data can, on the contrary, be too high to exploit the entire data set. Despite raising this question, authors do not provide any lead to a potential solution.

More recently, Le Goc et al. [11] introduce several notions useful to the conception of data physicalization: the notions of monolithic (composed of only one piece) and composite (composed of several pieces) data physicalizations. Composite physicalizations are associated to the notions of actuation, manipulability and granularity. Actuation corresponds to how autonomously pieces composing the data physicalization adapt to the context in which they are placed and to the data they represent. Manipulability corresponds to how easily pieces of a data physicalization are manually rearranged. Finally, granularity corresponds to the number of pieces the data physicalization is composed of.

Despite introducing these notions, they do not provide any guidelines concerning the conception or the design of data physicalizations. But they also introduce the notions of passive and dynamic

Why is this relevant?

the flow of this section needs work

shorten

too informal. Avoid

Even so, these

physicalizations which have become a subject of interest for researchers.

vague. Does not add.

3 CONCLUSION

In this review, one could notice that Data Physicalization is a field of research that raises many questions as well as the interest of many researchers. This review went through how the physicalization of

data could benefit users by enhancing the way they perceive and interact with data. Data physicalization can also be a simple and accessible way to understand and interpret data [12] and can be used in the presentation of said data to large-scale audience, problem solving or even comprehension enhancing. To help designers that get stuck with difficulties setting up these data physicalizations, some solutions have been developed and are already usable to accompany and support data physicalists. Researchers also take a great interest in providing design techniques for data physicalization, but research lacks experiments measuring the performance of such techniques. Finally, researchers tend to explore more and more the possibilities offered by dynamic data physicalizations which allow for continuous updates of the physical data representations. Future directions for this field of research could be about in-depth studying of design techniques performance for data physicalization. There is also a possibility to explore the enhancement of collaboration and interaction through dynamic data physicalization by bringing it into people's home. This would allow researchers to study the effect of such technology in real-life environments instead of imagining scenarios about it.

vague omit

suggest Not true remark

Weak conclusion

2.3 Dynamic data physicalization

Indeed, recent work shows that some researchers, while they emphasize the projects conducted about passive data physicalization, still find an insufficient amount of work has been granted to dynamic data physicalizations [11, 18].

Dynamic data physicalizations as opposed to passive ones allow for data physicalizations to rapidly and autonomously update when the data set requires it, to allow a continuous representation of data which can for example be useful to monitor data in real time, such as blood flow patterns [15]. To overcome this challenge, Le Goc et al. [11] introduce dynamic composite data physicalizations and propose a way of enhancing interaction, collaboration and decision making around data. For that they use wheeled micro-robots called *Zooids* which were developed in a previous project [10] during which they aimed to develop a new type of user interface: *swarm user interfaces*. *Swarm user interfaces* use a large number of autonomous robots that act both as *input* and *output* for the system so that they can be used for display and interaction at the same time. Using these *Zooids*, Le Goc et al. designed two user scenarios [11] in order to show how dynamic composite physicalizations could benefit collaboration and decision making. These scenarios describe how *Zooids* can be used to represent data in different ways, (e.g. automatically regrouping around a reference value from the data set according to each *Zooid*'s own value).

Dynamic data physicalization has also been studied by Taher et al. [18], which use the *EMERGE* dynamic visualization support to understand a presenter's gestures around the data representation, and their overall behaviour. Their research shows that dynamic physicalization help presenters to get a better grasp of their subject and eventually better expose it to others, while the audience shows a greater and more enthusiastic participation. An other part of this study has shown that participants were rather confident using the push-to-hide functionality of *EMERGE* bar charts. While most of dynamic data physicalizations are powered by touch-screen, this optimistic assessment demonstrates that non-mainstream controls physicalization would be also appreciated by both presenters and visualizers. Although results seem promising if one takes a look at the notions introduced by Le Goc et al., one might notice that *EMERGE* suffers from serious manipulability limitations, which is not the case for *Zooids* [10, 11] which can easily be independently manipulated.

A dynamic data physicalization can be obtained in different ways such as hybrid visualization which combines elements of physical representation with certain elements of virtual representation or by the presence of removable parts on the data physicalization. This type of representation can grant many advantages depending on the message that you want to convey. In addition, the use of dynamic physicalization of data can have social functions since it can allow a group to enhance communication among members [11, 18].

Don't start a section with

Remove redundancy

explain better otherwise we don't understand

unlike

avoid "you"

Not sure what you're saying

REFERENCES

- [1] Kathleen D. Ang, Faramarz F. Samavati, Samin Sabokrohiyeh, Julio Garcia, and Mohammed S. Elbaz. 2019. Physicalizing cardiac blood flow data via 3D printing. *Computers Graphics* 85 (2019), 42–54. <https://doi.org/10.1016/j.cag.2019.09.004>
- [2] JACOB BUUR, SARA SAID MOSLEH, and CHRISTINA FYHN. 2018. Physicalizations of Big Data in Ethnographic Context. *Ethnographic Praxis in Industry Conference Proceedings* 2018, 1 (2018), 86–103. <https://doi.org/10.1111/1559-8918.2018.01198> arXiv:<https://anthrosource.onlinelibrary.wiley.com/doi/pdf/10.1111/1559-8918.2018.01198>
- [3] Samuel Huron, Sheelagh Carpendale, Alice Thudt, Anthony Tang, and Michael Mauere. 2014. Constructive Visualization. In *ACM conference on Designing Interactive Systems in 2014*, ACM (Ed.). ACM, Vancouver, Canada. <https://hal.inria.fr/hal-00978437> We would like to thank Bon Adriel Aseniero and Lindsay MacDonald for their insightful comments.
- [4] Samuel Huron, Pauline Gourlet, Uta Hinrichs, Trevor Hogan, and Yvonne Jansen. 2017. Let's Get Physical: Promoting Data Physicalization in Workshop Formats. In *Conference on Designing Interactive Systems DIS 2017 (Proceedings of the 2017 Conference on Designing Interactive Systems)*. ACM, New York, United States, 1409 – 1422. <https://doi.org/10.1145/3064663.3064798>
- [5] Yvonne Jansen and Pierre Dragicevic. 2013. An Interaction Model for Visualizations Beyond The Desktop. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Nov. 2013), 2396 – 2405. <https://doi.org/10.1109/TVCG.2013.134>
- [6] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. 2013. Evaluating the Efficiency of Physical Visualizations. In *Proceedings of the 2013 Annual Conference on Human Factors in Computing Systems (CHI 2013)*. ACM, ACM, Paris, France, 2593–2602. <https://doi.org/10.1145/2470654.2481359>
- [7] Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. 2015. Opportunities and Challenges for Data Physicalization. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, New York, NY, United States. <https://doi.org/10.1145/2702123.2702180>
- [8] Yvonne Jansen and Kasper Hornbæk. 2016. A Psychophysical Investigation of Size as a Physical Variable. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan. 2016), 479 – 488. <https://doi.org/10.1109/TVCG.2015.2467951>
- [9] M. Karyda, D. Wilde, and M. G. Kjærsgaard. 2020. Narrative Physicalization: Supporting Interactive Engagement With Personal Data. *IEEE Computer Graphics and Applications* 41, 1 (2020), 74–86. <https://doi.org/10.1109/MCG.2020.3025078>
- [10] Mathieu Le Goc, Lawrence H Kim, Ali Parsaei, Jean-Daniel Fekete, Pierre Dragicevic, and Sean Follmer. 2016. Zooids: Building Blocks for Swarm User Interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST)*. Tokyo, Japan. <https://doi.org/10.1145/2984511.2984547>
- [11] Mathieu Le Goc, Charles Perin, Sean Follmer, Jean-Daniel Fekete, and Pierre Dragicevic. 2019. Dynamic Composite Data Physicalization Using Wheeled

- Micro-Robots. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan. 2019), 737–747. <https://doi.org/10.1109/TVCG.2018.2865159>
- [12] L. J. Perovich, S. A. Wylie, and R. Bongiovanni. 2021. Chemicals in the Creek: designing a situated data physicalization of open government data with the community. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 913–923. <https://doi.org/10.1109/TVCG.2020.3030472>
- [13] Renata G. Raidou, M. Eduard Gröller, and Hsiang-Yun Wu. 2020. Slice and Dice: A Physicalization Workflow for Anatomical Edutainment. *Computer Graphics Forum* 39, 7 (2020), 623–634. <https://doi.org/10.1111/cgf.14173> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14173>
- [14] Victoria; Esparza Antonio; Torres Rebeca; Napper Robbie Sosa, Ricardo; Gerrard. 2018. DATA OBJECTS: DESIGN PRINCIPLES FOR DATA PHYSICALISATION. In *DS 92: Proceedings of the DESIGN 2018 15th International Design Conference*. <https://doi.org/10.21278/idc.2018.0125>
- [15] Dolores Steinman, Peter Coppin, and David Steinman. 2017. Data physicalization for improving communication and memorization. In *16° Encontro Internacional de Arte e Tecnologia 16th International Meeting of Art and Technology*. 512.
- [16] Simon Stusak, A. Butz, and A. Tabard. 2018. Variables for Data Physicalization Units.
- [17] Saiganesh Swaminathan, Conglei Shi, Yvonne Jansen, Pierre Dragicevic, Lora Oehlberg, and Jean-Daniel Fekete. 2014. Supporting the Design and Fabrication of Physical Visualizations. In *Proceedings of the 2014 Annual Conference on Human Factors in Computing Systems (CHI 2014)*. ACM, Toronto, ON, Canada, 3845–3854. <https://doi.org/10.1145/2556288.2557310>
- [18] Faisal Taher, Yvonne Jansen, Jonathan Woodruff, John Hardy, Kasper Hornbæk, and Jason Alexander. 2017. Investigating the Use of a Dynamic Physical Bar Chart for Data Exploration and Presentation. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 451 – 460. <https://doi.org/10.1109/TVCG.2016.2598498>
- [19] Alice Thudt, Uta Hinrichs, Samuel Huron, and Sheelagh Cpendale. 2018. Self-Reflection and Personal Physicalization Construction. In *CHI'18: ACM SIGCHI Conference on Human Factors in Computing Systems*. ACM, Montréal, Canada. <https://doi.org/10.1145/3173574.3173728>
- [20] Wesley Willett, Yvonne Jansen, and Pierre Dragicevic. 2017. Embedded Data Representations. *IEEE Transactions on Visualization and Computer Graphics* (2017), 461 – 470. <https://doi.org/10.1109/TVCG.2016.2598608>