A Classification Schema for Data Physicalizations and a Carbon Footprint Physicalization

Steffen Haesler*
University of Würzburg

Jörn Hurtienne†
University of Würzburg

Franz Ertle‡
University of Würzburg

Patricia Theile§
University of Würzburg

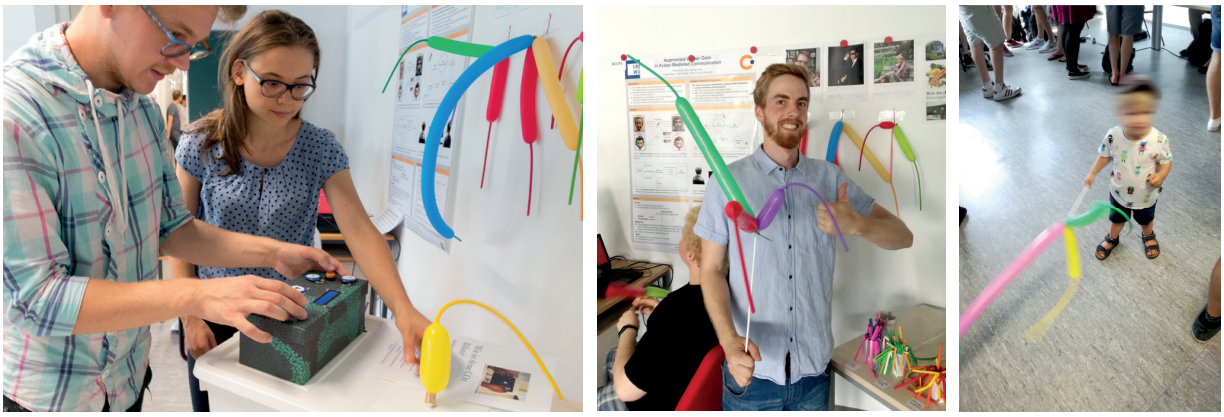

Figure 1: A high-fidelity prototype of a personal carbon footprint physicalsiation with twisting balloons.

Abstract

Based on previous literature, we propose a classification schema for data physicalizations and apply it in a case study to inspire the design of a carbon footprint machine. We present a high-fidelity interactive prototype that interactively physicalizes a user's carbon footprint with a balloon-based artifact. We describe the challenge to make different carbon footprints easily comparable. The prototype was shown and evaluated in a university-wide exhibition and the user study revealed the strength of the mapping. This case study shows that beyond the classification schema, issues of variable encoding remain significant and aesthetic, social and contextual factors need to be considered when aiming for a design language of data physicalizations.

Index Terms: Data Physicalization—Design—Carbon footprint—Sustainability;

1 A Data Physicalization Classification Schema

Can we find a classification schema for data physicalizations that not only helps us describe and understand the enormous space of existing data physicalizations, but that also helps us generating new ones? Hogan and Hornecker [1] presented a classification in which they analyzed more than 154 instances of data physicalizations, or as they called it, multi-sensory representations of data. During an open coding process they found several dimensions on which the data physicalizations could be analyzed. Among these were (1) the modalities of physicalizations (i.e. touch, sight, sound, smell, taste) and their combinations and (2) the choice of material for the physicalizations and whether these were related to the data. Other dimensions included (3) the purpose of the physicalization (i.e. casual, utilitarian), (4) whether live data was rep-resented, and (5) whether it was interactive. Their analyses revealed important design characteristics and opportunities for innovation. So, they found that most physicalizations used the modalities touch and sight, that also were the most frequent combinations (84%), while physicalizations using sound, taste and smell were very few in comparison. Physicalizations used a number of different materials, but a link between material choice and the data source was rarely found. About three-quarters of data physicalizations used archived (vs. live) data and only about two fifths allowed for interaction with the user. Three-quarters of the physicalizations were built for casual purposes, only a quarter was built for utilitarian purposes. Utilitarian physicalizations tended to make use of more modalities, a higher probability of using dynamic data, and were more interactive. Casual physicalizations tended to have fewer modalities, a higher probability of using static data and being non-interactive.

As this classification schema appeared to be quite useful, we used it as the basis of our own. We added new categories, changed the naming of some categories and features and left others away that could be more challenging to determine (e.g. emergent properties, the amount of insight a physicalization evokes). The final classi-fication schema is shown in Table 1, along with an example of the analysis of a data sculpture of the monthly average temperatures in Central Park NYC from 1869 through to 2015 [2].

We presented this classification schema to students who used it to analyze a number of data physicalization examples as an introduction to a data physicalization course [4]. The schema proved useful for getting a quick overview of the data physicalization design space. The schema was then used to guide the design of interactive data physicalization projects of the students. Here we discuss the example of a resulting data physicalization that represented the personal carbon footprints of its users. We describe the choice of data mappings, the actual prototype, the user interaction, and a user evaluation study. Both, the classification schema and the knowledge about underexplored areas in previous data physicalizations [1] influenced the design decisions along the way.

*e-mail: steffen.haesler@stud-mail.uni-wuerzburg.de
†e-mail: hurtienne@acm.org
‡e-mail: franz.ertle@stud-mail.uni-wuerzburg.de
§e-mail: irina_patricia.theile@stud-mail.uni-wuerzburg.de

Question	Possible values*	Example Central Park NYC	Carbon Footprint Physicalization
What is the original carrier of the data?	{data carrier}	Central Park in New York City	The user and his lifestyle affecting CO_2 emissions
Which and how many attributes are physicalized?	{attributes}	2 attributes: time and temperature	Weekly carbon footprint of a person in mobility, food and energy consumption
What are their scaling properties?	Categorical, Ordinal, Metric	Both metric	All metric
How many dimensions are used and how?	{dimensions and usage}	3 dimensions: time (months) on x-axis, temperature (°F) on y-axis, time (years) on z-axis)	2 dimensions: length of the balloon and time to inflate
What type of Visualisation?	Object or Attribute visualisation [3]	Attribute	Attribute
How many and which modalities are in the physicalization?	Sight, Sound, Touch (temperature, texture, firmness), Smell, Taste	Sight and touch (texture)	Sight: size and color of balloon. Touch: size and firmness of balloon. Sound: Annoying sound of the pump (the larger the carbon footprint the longer the sound)
What material? Is it linked to the data in some way?	{material}	White nylon plastic with a matte finish and slight grainy feel; no relation to data source	Gas in plastics; direct relation to data source
What is the intended purpose of the physicalization?	Work, Leisure, Education	Leisure	Leisure
Is the physicalization interactive?	Interactive, Non-interactive	Non-interactive	Interactive
Does the data physicalization use a dynamic data source?	Live data, Archived data	Archived data	Near-live data, becomes known in the moment of usage

*Possible values as specific value of an *enumeration* or as expression of a {*variable*}

Table 1: Data physicalization classification scheme based on a classification offered by Hogan and Hornecker [1] with example data of the Central Park NYC physicalization [2] and our carbon footprint physicalization.

2 A Carbon Footprint Physicalization

Data physicalizations are often static and non-interactive. By involving users in the process of creating physical artifacts from their own data, wanted to explore the more dynamic and interactive ends of the design space. In addition, we expected them to engage more with their data and reflect on the sustainability of their lifestyles. The result of the design process is a high-fidelity prototype based on an electric air pump that inflates twisting balloons proportionally to the carbon dioxide emissions of a user.

2.1 Mapping Carbon Footprints to Physicalizations

The choice of material for the physicalization should be linked to the kind of data that is physicalized. This is an important design gap in existing physicalizations [1]. Carbon dioxide emissions caused by the lifestyle of users are usually calculated as weights, e.g. kilograms. On a physical view, that makes sense, because the exact amount of gas molecules in a volume depends on air pressure and temperature. However, using weight for measuring gas has no correspondence in everyday experience. It is an abstract concept, and it is hard to imagine the dimensions of e.g. 500 kg of carbon dioxide. Also, the effects of carbon dioxide as an agent of climate change are tied to it being a gas present in the atmosphere. Preserving the gaseous aspect was important to us, so we thought about physicalizing carbon dioxide as a volume instead of weight. Human perception and estimation of volumes is, however, rather inaccurate and depends on the shape of the object and the possibility to physically feel the object [5]. One advantage of data physicalization is that it makes haptic interaction possible. But how would users be able to perceive relative differences in the amounts of carbon dioxide produced? First, we decided to show the volume of the gas via the size of its container. We tested different kinds of balloons and found that the perception of volume and volume change differed according to its shape. A normal round balloon that held 10% more volume than another did not look 10% larger. Indeed, the size difference was hardly perceivable. Still, while adding 10% more volume, an increase in size *was* perceivable, but it was difficult to get a sense of the exact amount of that increase. Volume change was markedly perceivable in the initial phase of inflating an empty balloon. Our tests ended with the insight that only twisting balloons, known for their use in creating balloon animals, have an ideal shape for perceiving volume changes. Inflating a twisting balloon primarily changes its length, which is much better perceivable than 3D volume changes because of the cubic function of the expansion in all axes. Using twisting balloons, different volumes of gas become comparable so that people can easily say if a balloon holds 30% or 10% more volume than another.

2.2 Carbon Footprint Data Sources

To transfer individual lifestyle information into carbon footprint data, we were researching applicable computation formats. We found scientific references for the following criteria: mobility [6], food [7], and energy consumption related to household size [8]. Then we scaled the output of these formulas to carbon dioxide equivalents per week ($kgCO_2e$/week) because we deemed it easy for participants to give input based on a typical week. So, for example, a vegetarian diet leads to 27 $kgCO_2e$/week compared to 40 $kgCO_2e$/week for a meat eater. To keep the interaction simple, we made some assumptions like a consumption of 2000 calories of food per day, that households have a washing machine, or that a car has a medium size. In our scenario, users needed to give the following input to estimate their carbon footprint:

1. *Food:* Vegan, vegetarian, meat.
2. *Mobility:* (km per week for each method) Bike, bus, tram, regional train, long-distance train, car.
3. *Energy consumption:* Electricity: Renewable or conventional energy. Household size: '1' to '5 or more' persons.

2.3 Mechanism for inflating balloons

To control the amount of gas used to inflate the balloons we considered different kinds of pumps. Classical balloon pumps were not suitable because they would let the air escape as soon as they were turned off. So we used a 220V electric air pump and modified the vent with a hose and an adapter to fit a balloon. To control the pump, we took an Arduino Nano micro-controller that switched a relay that powered the pump. First of all, we had to calibrate the filling of a balloon, so we measured the amount of time the pump needs to fully inflate a balloon. Then we mapped a full balloon to a certain amount of carbon dioxide, which we decided should be equivalent to 1000 kilometer driving a medium sized car a week or 168 kgCO_2e. Then we were able to map the amount of carbon dioxide to seconds activating the pump. The results of this calibration were implemented in the C-based code of the Arduino.

2.4 Interaction

We explored different input mechanisms for choosing the categories and entering values. To keep within the paradigm of controlling a mechanic or pneumatic machine, we refrained from using touchscreen or keyboard and mouse input. Instead, we employed mechanical buttons, knobs, and a small LC display capable of showing two lines of text. To change categories we used RFID chips (Figure 3). Altogether, the interaction had six basic steps:

1. Take an RFID chip with the category symbol on it.
2. Put the corresponding balloon on the vent.
3. Put the RFID chip on the scanning area and read the question on the display.
4. Enter the related amount or answer with a rotary control.
5. Confirm the answer and see how the balloon inflates.
6. Change the category.

Using one balloon for the combined volume of all categories we felt it was difficult to compare the contribution of the different categories to the overall carbon footprint. During the inflation process one could hear the pump working and see the balloon growing, but the finished balloon gave no indication of how much the resulting gas volume was due to the user's diet, mobility or energy consumption. To allow for easier comparison we used three balloons per user with different colors that represented the three lifestyle categories. Now it was not only possible to compare the carbon footprint between the categories of one user, but also to compare the footprints between different users. To facilitate this we also devised three personas that were described on a small poster and their respective balloons hung beneath them (Figure 1, Figure 2).

2.5 Portability

Balloons are often put on a stick handed to children at neighborhood events. Thus, we also wanted a portable solution for our balloons and wanted to place our machine in a public space. So we bought balloon sticks and mounted the three balloons of one user's session on top of each stick. We presented the carbon footprint machine at a university-wide exhibition and hoped that the balloons would cause curiosity when other people were walking around with their balloons. We also discussed whether we needed to invert the inflation process because with the envisioned design people with a small carbon foot-print would have only partly inflated balloons. It could be that they felt they had less aesthetically-pleasing balloons than people with larger balloons (and larger carbon footprints). As we

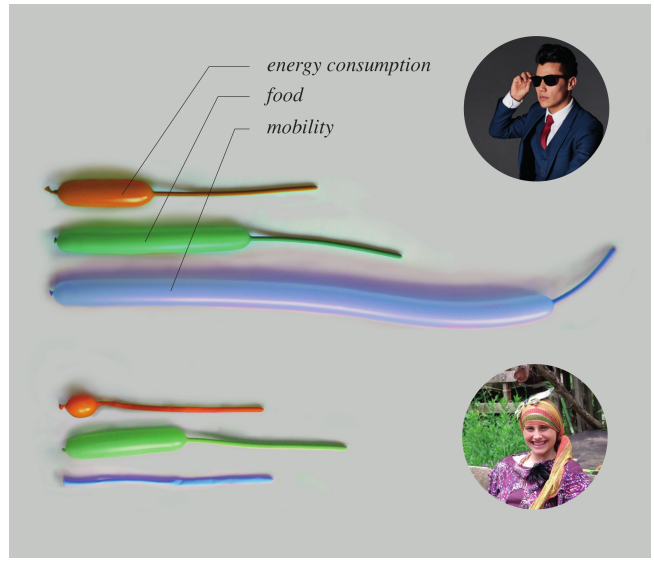

Figure 2: Personas demonstrated typical results for comparison. The top shows the less sustainable lifestyle of a person living alone using conventional energy (30 kgCO_2e), eating meat(40 kgCO_2e) and driving 600 km per week by car(100 kgCO_2e). Below is a very sustainable lifestyle living in a shared apartment with four people using ecological energy(1.5 kgCO_2e), having a vegan diet(20 kgCO_2e) and getting around by bicycle (0 kgCO_2e).

did not know how people would react, we decided against a redesign of the original idea, as this would have eliminated the easily comprehensible mapping of more emissions to more volume. Rather we conducted a user study at the exhibition to find out about the effectiveness of our approach.

3 User study

We presented the prototype at a local university exhibition where visitors were able to use it and we asked several of them to participate in a qualitative interview. The qualitative interview was done based on a previously developed guide using methods of Kruse [9].

Fourteen visitors (four female) agreed to participate. Eleven were students, two were academic staff, and one was an external visitor. Their age ranged between 20 and 55 years; most participants were within the range of 20-25 years. The carbon footprint machine stood on a table surrounded by the posters of personas with their carbon footprint physicalized in balloons. Other posters displayed general information about climate change and a short text about how to interact with the prototype. Users took the RFID chips and three balloons and used the machine as explained above (Figure 3). They pulled a balloon of a certain color (Figure 2, Mobility: Blue or Purple, Food: Green, Energy consumption: Red or Orange) on the top of the vent and entered their lifestyle variables using a rotary control and confirmed their selection with a button. Then the equivalent amount in CO_2KGe/week was shown on the display, the pump started very noisily to inflate the balloon and the display showed the progress of inflation.

After watching (and hearing) the inflating process, they took off the balloon and strung it on top of a stick and proceeded with the next RFID chip to do the same with the next category. In the end, users could take away with them their personal carbon footprint physicalized in the three balloons, and compare their result with the balloons of the displayed personas or the results of other visitors. Before and after the interaction with the prototype, we asked the interview questions.

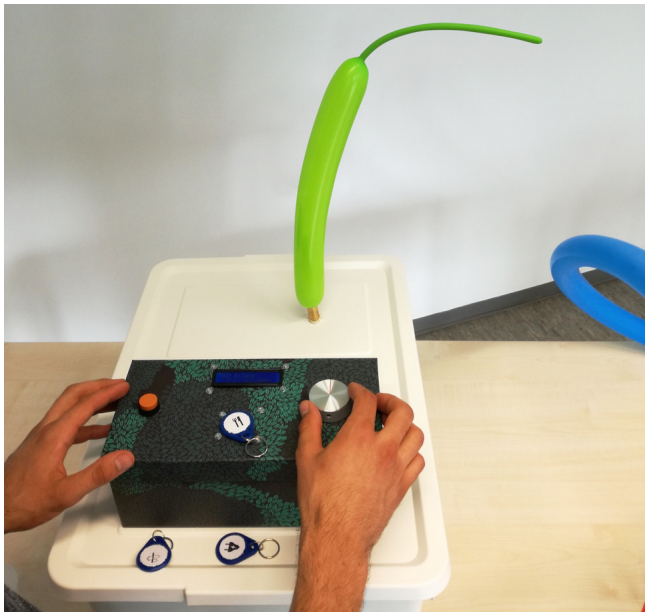

Figure 3: The user enters data about her diet and then pushes a button to get the physicalized equivalent of $kgCO_2$ in shape of a balloon.

The results of the user study revealed that none of the 14 participants had thought in detail about their carbon footprint before. After seeing their results, users were surprised of the difference between the categories. Six participants were surprised that their carbon footprint was higher than expected; three had expected a higher amount; and five participants did either expect their results or had no prior opinion. Nine participants were motivated or highly motivated to change behavior after using the prototype; two felt better informed and another two did not feel motivated at all. Regarding usability, the participants remarked that they had difficulty at the beginning to find out how the interaction with the RFID chips worked and that the rotary control was not behaving on an intuitive scale. Asked about the physicalization with twisting balloons all of the participants had positive thoughts about it and mostly claimed that the idea was "cool", "funny", "cute" or "super". Three participants explicitly mentioned that they found the mapping to a gas very suitable, while one person thought that the mapping was too obvious. Only one person regretted that a smaller footprint leads to a less inflated balloon, while others were not concerned about this. However, nearly all participants had mixed feelings about the balloons. On the one hand, they were proud of the personally meaningful physical object to take away. On the other hand, they sometimes felt bad about being reminded of their carbon footprint and about visibly sharing their data with other people. Regarding the color choice of the categories, users saw room for improvement. While green balloons for food and yellow/red balloons for energy consumption were accepted widely, users wanted to have blue balloons for mobility instead of purple. In addition to the given categories, participants suggested having other balloons, e.g. showing a base value as part of the society in a certain country. They considered to add heating to energy consumption, to add a question whether food is sourced locally or not, and to include travel behavior and other spare time activities.

4 Discussion and Conclusion

The design choices made when building the carbon footprint machine touch upon different aspects that are relevant when considering design languages for data physicalizations. First, we were led by the classification schema presented in the introduction (Table 1). Throughout the design we looked at areas that were under-explored by previous data physicalizations. We wanted to have an interactive prototype that flexibly could adapt to each user's data. We wanted to use a meaningful material reflecting what the data is about. Through the prototype we coded the carbon footprint data in several physical dimensions and modalities. The first dimension is a process parameter: the time to inflate the balloon (that was accompanied by the noise of the air pump and certainly drew the attention of other visitors nearby); the second dimension is its physical result: the size of the balloon that was easily perceivable via sight and touch. Thus, the classification schema allowed us to become aware of how to locate our prototype within the design space (Table 1) and take decisions on how to develop it into interesting directions. Second, we encountered the classical issues of mapping the data to physical parameters and found a way to physicalize volume changes (that are not easily perceived by people in 3D) in a way that made volume change perceivable along one dimension. Along with others [10] we see the necessity to derive a design language for mapping data to physical parameters of which there are more and more diverse than in the classical case of 2D visualizations [11]. Third, we also need to go beyond the encoding of variables in physical dimensions [12] and look at the context in which physicalizations are used and the specific denotations and connotations that the chosen material holds. The noise of the pump and the colorful bundle of balloons had aesthetical expression and triggered social reactions, i.e. it became meaningful beyond the actual data mappings. A design language of data physicalizations would also need to tell us about the aesthetic and the social, as well as the contextual effects particular design decisions imply. This will be by far the most difficult task (if it is at all possible), but the one that might be worthwhile because it could tell us more about what is so unique about physicalization beyond classical visualization.

References

[1] T. Hogan and E. Hornecker, "Towards a design space for multisensory data representation," *Interacting with Computers*, vol. 29, no. 2, pp. 147–167, 2017.

[2] "2016 – central park nyc temperatures." `http://dataphys.org/list/central-park-nyc-temperatures/`, accessed 2018-07-03.

[3] R. Spence, *Information visualization*, vol. 1. Springer, 2001.

[4] J. Hurtienne and D. Reinhardt, "Teaching data physicalisation to hci students - a case report," in *Designing Interactive Systems Conference (DIS '17). Edinburgh, UK*, (Edinburgh, UK), 2017.

[5] M. Kahrimanovic, W. M. B. Tiest, and A. M. Kappers, "Seeing and feeling volumes: The influence of shape on volume perception," *Acta psychologica*, vol. 134, no. 3, pp. 385–390, 2010.

[6] R. Grießhammer, E. Brommer, M. Gattermann, S. Grether, M. Krüger, J. Teufel, W. Zimmer, and R. geförderten Projekts, "C02-einsparpotenziale für verbraucher," *Freiburg: Öko-Institut eV Online: http://www. oeko. de/oekodoc/1029/2010-081-de. pdf (29.04. 2013)*, 2010.

[7] P. Scarborough, P. N. Appleby, A. Mizdrak, A. D. Briggs, R. C. Travis, K. E. Bradbury, and T. J. Key, "Dietary greenhouse gas emissions of meat-eaters, fish-eaters, vegetarians and vegans in the uk," *Climatic change*, vol. 125, no. 2, pp. 179–192, 2014.

[8] S. Kleinhückelkotten, H. Neitzke, S. Moser, *et al.*, "Repräsentative erhebung von pro-kopf-verbräuchen natürlicher ressourcen in deutschland (nach bevölkerungsgruppen)," 2016.

[9] J. Kruse, *Qualitative Interviewforschung*. 2015.

[10] Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk, "Opportunities and challenges for data physicalization," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3227–3236, ACM, 2015.

[11] J. Bertin, "Semiology of graphics: diagrams, networks, maps," 1983.

[12] C. Ziemkiewicz and R. Kosara, "Beyond bertin: Seeing the forest despite the trees," *IEEE computer graphics and applications*, vol. 30, no. 5, pp. 7–11, 2010.