Teaching Data Physicalisation to HCI Students – A Case Report

Jörn Hurtienne
University of Würzburg
97074 Würzburg, Germany
hurtienne@acm.org

Daniel Reinhardt
University of Würzburg
97074 Würzburg, Germany
daniel.reinhardt@uni-wuerzburg.de

Abstract

Data Physicalisation (DataPhys) is a new area of research at the intersection of Information Visualisation and Tangible Interaction. When designing a course on DataPhys for HCI students, neither a teachable canon nor pre-configured course material has been available. Here we share our experiences with conducting a course on DataPhys and discuss any lessons learnt.

Author Keywords

Data Physicalisation; Information Visualisation

Teaching Data Physicalisation to HCI Students

Although data physicalisations have been around for several 1000 years, their systematic study has only recently begun. Although it is too early to expect a teachable canon of the area, the intersection of information visualisation (InfoVis) and tangible user interfaces (TUI) still makes for a rewarding area to teach in HCI. If posed as a research course, students have the ability to learn basic concepts, apply these in a practical project of their own, and even explore yet-uncharted territory in data physicalisation research. Here we give impressions from running a course on data physicalisation that we held as an elective for HCI students at the University of Würzburg, Germany. The course took place during the summer semester in 2016. It ran on a biweekly schedule with seven four-hour sessions. The sessions were divided into three parts: Introduction, Special Challenges and Project Work.

In the *Introduction* part, students learned basic concepts and purposes of information visualisation; they saw historic and current examples of InfoVis and discussed the value of digital visualisations (e.g. easy re-ordering, filtering, "brushing") [3]. InfoVis then was contrasted with data physicalisation and its possible benefits (e.g., better use of active and multimodal perception, higher accessibility, fostering data understanding and viewer involvement) [2]. Special emphasis was given on discussing the differences and commonalities between tangible interfaces and data physicalisations. Then, to deeper understand the design space and nature of previous data physicalisations, a framework of analysis was presented and students explored different examples of data physicalisations using the online repository at dataphys.org/list. The results could be compared to the more thorough analysis of [1] to give a bigger picture on the design

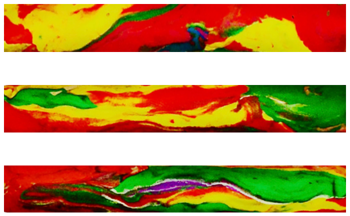

Figure 1: Personalised data physicalisation: each stick represents a person. The colours encode the importance of different variables over the lifespan (the length of the stick). These objects can be individually identified by their colour patterns and support rich multivariate comparisons.

Figure 2: Personalised data indicated by form and colour. Red elements (on top of cars) code age by size; green elements code gender by form; yellow elements indicate whether one is born in the city; blue elements code religious confession by form.

space and opportunities for further exploration. Finally, as the first hands-on exercise in producing data physicalisations, students physicalised a simple data set themselves (e.g., car ownership across different European countries, the development of unemployment numbers in Germany, production of red and white wine in the region). This exercise was done in class where a large range of materials, e.g. Lego bricks, marbles, toy cars, even small robots (OzoBots), was available to tinker with. The results of the exercise were discussed in class.

The *Special Challenges* part of the course aimed at exploring areas in data physicalisation that are currently underdeveloped (and, possibly, difficult to realise). First, most current data physicalisations present only a few data objects (e.g., countries) and variables (e.g., GDP) at the same time. Most of the time, these representations are static and non-interactive. Many datasets in real life, however, are mass data, e.g. from demographic surveys involving several thousands of respondents including data on a large range of variables. How can we enable researchers or even the general public to explore such large data spaces? To introduce the problem to the students, a use case was presented, in which an ergonomics researcher wanted to explore the influence of computer use and mobile work on working conditions and the experience of stress at work. The data set came from a representative social survey with over 20,000 respondents. The classical approach would be to use statistical software to filter data, define variables and run routines for the descriptive and inferential statistical analyses – a very cumbersome process in which the results are presented as intangible numerical values. The challenge was to make this research process more tangible by physicalising the data and finding solutions for selecting, filtering, combining and separating physical data points. The students needed to address many questions at once: How should data be represented? How could users specify queries? How does the system respond and guide the user around the inevitable difficulties present in the physicalisation? To spark the imagination, a data physicalisation with rice (Stan's Cafe: Of All the People in All the World, see www.youtube.com/watch?v=iDWcuBygAUw) was presented and discussed in class before students developed their own concepts, again using a wide range of materials.

Second, mass data are often about people. The second challenge, therefore, posed the question of how data physicalisations can retain or emphasise the individuality of each person that defies simple categorisations, is shaped by a rich history and may even resist close scrutiny in public. Some attempts at "giving statistics a face" can be found on YouTube. For example: "If the world today were shrunk to the size of a village of just 100 people…" then 7 would speak English, 53 would be of Asian origin, 10 would have control over nuclear weapons and so on (e.g. www.youtube.com/watch?v=jNnbO8x4JAY). Although making for compelling visualisations, the relations between the single characteristics of these people do not become clear (e.g., how many and which of the Asian people have access to nuclear weapons?). This problem was elegantly solved in a data embodiment performance by theatre producers Rimini Protokoll who brought 100 people of one city on a theatre stage. These people were selected to be representative of their city (according to age, gender, marital status, ethnicity, and neighbourhood). They responded to

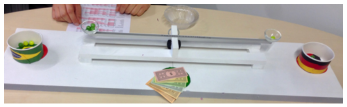

Figure 3: Country Balance. The user chooses countries and the topic to be compared (here: per capita income of Brazil and Germany). Then marbles can be added to both sides of the balance. When balanced, the income ratio between the countries can be seen immediately. Physicalising different ratios is possible by moving the beam across the fulcrum according to the user's selections.

Figure 4: Deforestation of the Amazon Rain Forest. Users could earn money by physically removing trees and laying out pastures, soy bean plantations and streets. Red LEDs show the area of rain forest that was cleared during a user-chosen time period. A map of Germany is included to impart a sense of the involved areas.

about 100 questions, not unlike those in a social survey, with different forms of embodiment. They, for example, would walk to different areas on stage labelled *Me* or *Not Me*; they would hold up colour-coded labels, raise their hands and so on. Students analysed videos of these performances in different cities (e.g. Amsterdam:www.youtube.com/watch?v=lHyTCBnqTbc, Berlin: vimeo.com/40925638, Melbourne: vimeo.com/49825849). Then they discussed in class the choices Rimini Protokoll made to visualise mass data and the techniques that they employed to make the data personally meaningful. Students subsequently worked on the problem of how these personalisation techniques could be transferred to data physicalisation tasks. Again, they physicalised their ideas with a range of different materials. Two examples are shown in Figures 1 and 2.

In the *Project Work*, the emphasis was on building a well-developed concept of a data physicalisation and to present this concept as a low-fidelity prototype. Low fidelity, for example, means that the final material, if difficult to realise, did not need to be present in the prototype or that the interaction of users with the data could be facilitated by a Wizard of Oz. The project documentation should make clear the choices made during the project: which data was physicalised in what form and with which material as well as how the data physicalisation was made interactive. To find a relevant project, several ideas were pitched and the most interesting and feasible were selected for further work. Students were completely free on which data to visualise, but were told that extra points could be gained when taking on the challenges of mass data physicalisation and personalisation. Several sessions were spent working on the progress of each project by offering design critique and discussing possible alternatives for the realisation of the projects. The prototypes were presented at the final session of the course and the work was documented in written reports. Three projects were realised: Country Balance (Fig. 3), Deforestation of the Amazon Rain Forest (Fig. 4), and Playing with the Power Grid of Germany (Fig. 5).

Lessons Learnt and Future Work

Altogether, the course has proven effective in introducing HCI students to basic concepts in information visualisation and data physicalisation and involving them in conceptualising data physicalisations that venture into yet-uncharted territory (mass data and personalisation). Their physicalisation projects equipped them with new manual skills and fostered their interests in tangible and physical computing. Although it might be too early to draw final conclusions, after this first implementation of the course several lessons can be taken away to inform the next iteration.

First, the *Introduction* proved worthwhile, as the contrast between traditional InfoVis and data physicalisation was nicely worked out. The classification exercise introduced students to the variety of available physicalisations and their characteristics. So, students were sensitised to the design space and to specific questions of material selection, modality of representation, interactivity and purpose of the physicalisations. With a view on the later presented challenges in the second part of the course, this exercise also solved to ground the students and remove anxieties: most successful work in data physicalisation is still non-interactive, mainly in the visual modality, not very complex with regard to the number of data objects and variables and does not care much about

Figure 5: Playing with the Power Grid of Germany. Across a map of Germany, several high-voltage power lines are drawn. Users can place wooden tokens, representing power plants, on the map. Depending on their size and type, these power plants can provide energy to specific areas on the map that light up when sufficient capacity is reached. Electric arcs are sparked when producing more energy than needed. Overcapacity can be distributed to other areas by manipulating the coloured caps on top of power-line transmission nodes. The aim is to simulate the planning of power grids and takes into account the costs and earnings of constructing power plants and transmission lines as well as the energy supply that is possible due to geographic features (e.g. wind energy at the sea or in the mountains).

material selection [1]. It, however, also showed where the opportunities for further work lie. The first exercise of physicalizing a simple data set worked well, especially with the materials provided. Going beyond simple material at this stage, e.g. providing technologically more advanced artefacts like OzoBots to show actuation capabilities only served to take attention away from the actual physicalisation issues. Instead students liked to play with the robot and test its capabilities and constraints. To avoid technological distraction, in the future we therefore consider introducing a separate session in which novel actuation and fabrication technologies are presented and can be tinkered with.

Second, the *Special Challenges* were indeed challenging to our students. They did well in trying to understand the problems presented and finding solutions. When presenting the mass data challenge, the use case will need to be introduced more transparently to ensure a clear understanding of the task. Solving the problem proved to be complex, because the physicalisation and the interaction requirements (filtering, selecting etc.) had to be addressed at the same time. In the future, we would try to address the different requirements one by one with each group of students working on a single problem at a time and later combining all the results. The personalisation session went better, partly because there was a fixed structure set by the analysis of the data embodiment of Rimini Protokoll. The first prototypes included some excellent ideas (Fig. 1), but more formal ways of personalising prevailed, e.g. abstract colour and form coding of people's characteristics that relied on first reading a legend to understand the codes (Fig. 2). For a first prototype we found this ok, given that our students are mainly trained in computer science and psychology, lacking a distinctive training in design.

Third, the *Project Work* started with a pitch of project ideas that was prepared quite informally and therefore sometimes missed systematically looking at originality/feasibility criteria of evaluating the single concepts before the pitch. Better preparation would be necessary to avoid this. The data physicalisation concepts that emerged in the project phase were quite interesting and showed that a good concept can be easy to realise as a prototype (Fig. 3) or entail a lot of manual work (Fig. 4) and that a clear concept of the game mechanics is needed for a physical data game (Fig. 5). In the future, we would make the presentation of prototypes available to a larger audience joining a general term-wide exhibition of all HCI courses.

Fourth, in future iterations of the course, we want to move closer to implementation, i.e. introducing and using physical computing (Arduino) or 3D fabrication techniques to facilitate more advanced versions of the prototypes. Also, we need to find ways of scaling the teaching concept from now 6…8 participants to the usual size of classes that are more in the range of 15…20 participants. Although we may focus more on simpler data physicalisations, the challenges will remain to help push the boundaries of the field. The challenges also provided a springboard for more intensive work on these topics (a paper and a Master's thesis derived from these).

Please contact the authors with any thoughts on and experiences in teaching DataPhys to HCI students!

References

1. Hogan, T. et al. (2017). Towards a Design Space for Multisensory Data Representation. *Interacting with Computers*, *29*(2), 147-167.
2. Jansen, Y. et al. (2015). Opportunities and challenges for data physicalization. In Proc. of CHI ’15, ACM, 3227-3236.
3. Spence, R. (2014). *Information visualisation: An introduction*. Cham: Springer.