


Voyage Viewer: Empowering human mobility at a global scale

Isabella Loaiza¹ , Tobin South¹ , German Sanchez², Serena Chan³, Alice Yu¹, Felipe Montes², Mohsen Bahrami⁴ ,
Alex Pentland^{1,4} 

¹MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

²Social and Health Complexity Center, Universidad de Los Andes, Bogotá, Colombia

³Department of Political Science, Wellesley College, Wellesley, MA, USA

⁴MIT Connection Science, Institute for Data, Systems and Society, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

The challenge of refugee relocation is fertile ground to pose a new direction in the quest for extended human intelligence: developing systems that leverage big data, and the power of social learning to provide personalized visual analytics for big life decisions. To probe into this new avenue, this paper presents Voyage Viewer, a novel open-access multi-stream data dashboard called Voyage Viewer. It helps individuals make their own relocation and migration decisions given personalized queries and visualizations, which stands in contrast to previous top-down approaches that use algorithms to match individuals and places, as is the case for some refugee relocation programs. Voyage Viewer hopes to foster social learning between community members to improve the match between migrants and their potential new communities so that both can reap the benefits of the move.

CCS Concepts

• **Human-centered computing** → **Visualization toolkits**;

1. Introduction

Relocation and migration decisions continue to be a complex personal challenge for many. Assessing the right fit between individuals and migration destinations requires insights derived from others with similar preferences, skills, values and histories. Indeed, the need for personalization is exemplified by the importance placed by organizations like Netflix, Amazon, and Spotify on capitalizing personal data to provide tailored recommendations.

Many of these recommender systems target higher-volume, lower-value decisions [DED19], but close to none target big life decisions. When deciding where to go to school, what career to choose, or whether to relocate to a new city, most wrestle with these decisions in the same way our predecessors did: with anecdotal evidence and little or no help from AI and Big Data.

This paper seeks to shed light on a new research direction: developing systems that leverage big data, and the power of social learning to provide personalized information and visual analytics (VA) to help individuals make big life decisions. The nature of transformative, high-cost decisions is different from low-cost decisions. Their feedback is often fuzzy, delayed, and not as easily quantifiable by metrics like skipped songs or watch time. People tend to make fewer big decisions in their lifetimes, making fewer data points available per person and increasing the system's need to use social exploration to derive insights. Thus, this problem cannot be reduced to the application of existing algorithms and VA methods used to recommend movies or purchases to careers and

migration. Systems in the vein that we propose do not seek to automate human decision-making but rather bolster agency in helping people make their own informed decisions, not have an AI make choices for them.

We think the challenge of refugee relocation is fertile ground to test this new direction while helping alleviate some of the burdens endured by migrant and refugee communities. The refugee relocation algorithms that rose to popularity during the Syrian refugee crisis were a stepping stone in this direction. These algorithms helped countries cope with the influx of newcomers, but had important limitations like the choice of variables to make the match [CBG21, BFH*18, KCB21], and that they were used to mostly replace, rather than aid human decision-makers. They were criticized because of the complex power relations they involved and their part in further eroding the agency of a vulnerable population, among an array of other ethical concerns [BHCM19, MK20, DB15].

To probe into this new avenue, we implemented in a public data dashboard called Voyage Viewer. Voyage Viewer aims to foster social learning between community members to improve the match between migrants and their potential new communities so that both can reap the benefits of the move. Voyage Viewer embraces a decentralized, grassroots approach where power and decision-making are in the hands of many instead of a selected few, creating self-governed communities that can learn and are empowered to make informed decisions. This decentralized paradigm makes the resulting communities more resilient and reduces unintended consequences [PLH21, HP19].

In summary, the main contributions in this paper are:

- Posing a new research avenue about personalized VA for big life decisions.
- Designing a VA tool to help decision-making at an individual level and return some agency to disenfranchised communities like refugees and asylum seekers.
- Fostering a better matching process between potential migrants and destination communities.

2. Related Work

Several tools exist to visualize the flow of migrants and refugees [SO17, CLS, BCG*17, Ije14, SAA*21], humanitarian emergencies [ftCoHA, fRU16], and the impact of border controls [Spi07, MIC20, WP11], and complex citizen displacements [SLM21]. These visualizations stem from a broader history of tools to visualize and understand several aspects of the social world like economic development [SH11], the flow of goods [Com17], city life [MCP19, MCDP21] and local economic opportunities [BAS*20].

However, Voyage Viewer is the first to use user-provided input to produce personalized visualizations to help potential migrants asses the match between their needs and potential destinations. In this regard Voyage Viewer is more similar to other works that focus their attention on how technology helps migrants in their journeys [CCCS19].

In its use of visualization methods, this paper connects to a broader body of work about the technical aspects of visualizing flows [Tob87, WDS10, AAFW16, Guo07]. Similarly, challenges to multi-scaling space and time have been detected before, visualized with different scales of spatial representation, which aid users in gathering information about details of distance by zooming into the data and relations between areas of migration flows [AAB*13, SLM21]. Voyage Viewer would stand as a complement to these technologies, not as a replacement, because it can provide additional data from multiple sources as a dashboard and transforms it into consumable information [Smi13].

3. Design Rationale

Our design was based on the three cornerstones of the design triangle framework by [MA14]: Users, Data, and Tasks.

- **Users.** The main target audience for Voyage Viewer are individuals. Users are not circumscribed to a particular geography because in the future we hope to have data from as many countries as possible. Similarly, the dashboard is suitable for a wide range of computer skills from experts to users with average computer literacy and internet access. The tool is especially useful for people interested in relocating, international migrants, or those who are curious about potential opportunities outside their current place of residence.
- **Data.** To visualize destinations, employment status and income distributions Voyage Viewer uses government household surveys between 2012-2020 like the American Community Survey (ACS) [Bur] and the Encuesta Integrada de Hogares (GEIH) [dED]. For social indicators like the tone of news about migrants, public data from the GDELT project was used [GDE].

NUMBEO provided data about the cost of living in cities across the world [NUM]. Data from the United Nations High Commissioner for Refugees (UNHCR), the International Organization for Migrants (OIM), Deaths at the Borders Database [Las15], and the Australian Border Deaths database [MIC20] are used in other sections of Voyage Viewer.

Main challenges associated with data were the standardization of country labels across different datasets, marrying multiple data streams to tell a cohesive story, and maintaining homogeneity in the information available across countries given disparity in government data availability.

- **Tasks.** Citizens solved tasks developed with the help of relevant literature [Bar10] as well as inspired by the experiences of the international migrant team behind Voyage Viewer. The tasks are as follows: 1) identifying the main national and international destinations for people similar to them 2) Assessing their potential job prospects, levels of income, and cost of living at frequent destinations for people like them 3) Describing the impact of the information presented in Voyage Viewer in a hypothetical migration scenario.

4. Design Implementation

The Voyage Viewer section discussed in this paper is the *People like me* page. This page is meant to help citizens and migrants learn from other migrants' decisions and outcomes to make informed decisions about the cities or countries to move to. We also include data about locals to help users asses the gaps across economic and social indicators between locals and migrants. The page's user flow is show in Figure 1.

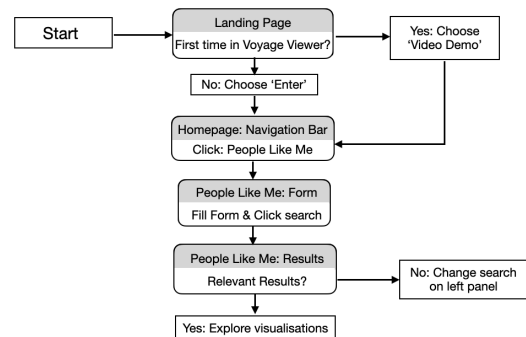


Figure 1: Figure 1 shows the user flow for the *People Like Me* page on Voyage Viewer.

The page has the following sections: Initial form, and Main page, which is in turn divided into four parts comprising an array of indicators: number of similar people, national and international destinations, and economic outcomes, and social indicators. By allowing users to input the parameters they want to be matched on, Voyage Viewer's approach is more transparent and easily interpretable than other mathematically driven similarity measures..

To access the results page with the visualizations, users fill out a standard form to determine similarity to others in Voyage Viewer's data. A description of the fields and a picture of the data are included in the SI.

After the initial form, the People Like Me page will appear [Figure 2](#). It consists of a main section with data visualizations (B) and a left panel (C) to display or adjust the input parameters from the initial form. The lower part of the left panel shows the data sources used to create the visualizations. The main container is divided into 3 subsections. The first subsection is the number of ‘People like me’ (D). This section shows the number of people in the data that match the user input variables given in the initial form. This is included for transparency, so that users know the visualizations are created few or many examples. The destinations section (D) displays the most frequent national and international destinations that match the user input in two separate treemaps. The treemaps display the names of the cities or countries for national and international destinations, respectively. The treemaps have tooltips that show the percentage of people at each destination for more clarity.

The economic outcomes section is the most visually diverse. It contains a heatmap [Figure 3a](#), density plots, and several vertical bar charts. The heatmap shows the percentage of people who match the user input and are employed, unemployed, or inactive in the top five most frequent national destinations shown in the section above. Similar to the number of people like me, the heatmap has an extra column that shows the total number of people in each region for transparency. Percentages can be misleading, and having 100% employment in a region where there are less than 20 migrants is very different from a region where 95% of the 100,000 migrants that live there are employed. To further aid in the comprehension, clickable dropdown definitions from the U.S. Bureau of Labor are included. The density plots show the migrants’ income distribution in the top 3 national destinations, and are overlaid by the locals’ income distribution [Figure 3b](#). This helps users estimate their level of income but also estimate the gap between local’s income and migrant’s income. Previous research has shown that users can effectively estimate quantities from density plots over other types of plots and are considered visually appealing [KKHM16]. Finally, bar charts show the cost of living in the top national and international destinations for people similar to the user input (Images in the SI). The employment section describes the most frequent ways to get a job for migrants [Figure 3c](#) that match the user input compared to locals at destinations.

The social indicators section shows line plots that depict the temporal changes in the tone of news about migrants in the national and international destinations for migrants that match the parameter input by the user. This can provide an idea of how welcoming is the discourse about migrants in the media in potential destinations.

5. Evaluation

5.1. Evaluation Procedure

For evaluation of the VA design we conducted a usability evaluation and a third-party accessibility evaluation. The usability evaluation included a total of ten participants, who were all located in Colombia, and a mix of genders. Participants generally commented favorably on the website, especially the ability to find, compare, and analyze information. The usability evaluation included three tasks, time for free information exploration on the website, and unstructured feedback about general impressions. The usability study gathered anonymous information about age range, gender, educational

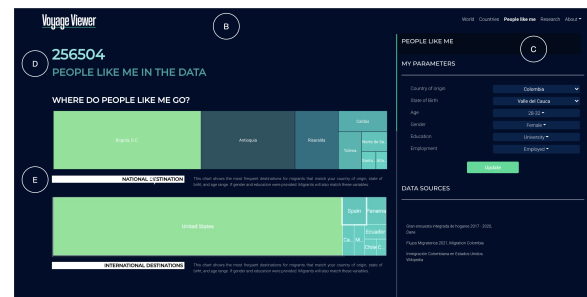


Figure 2: A. Figure shows the People Like Me main page. In the image the different sections are labeled with letters to show the different parts. B. The main container with all the visualizations. C. Left panel to display or change user input to build the visualizations. On the bottom part of the panel, we display the data sources used to make the visualizations. D. Shows the number of matches found in the data based on the user input. E. National and International destination treemaps. Tooltips showing the exact fraction of matches that go to any given destination appear upon hovering in each box.

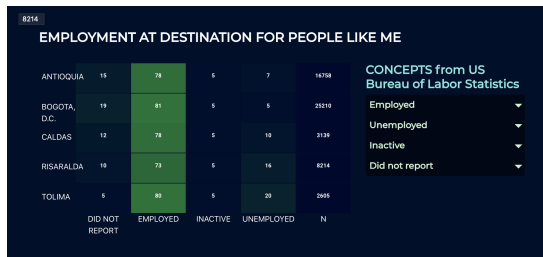
background, and whether the participants had migrated nationally or internationally for longer than two months. We showed a short video about the tool’s overall goal and potential uses to contextualize users and ensure that all participants were exposed to the same information. The following evaluation phase presented users with a hypothetical scenario where they were put in the shoes of a potential migrant. They were asked to assume the role of an unemployed female between the ages of 28 to 32 from Santander, Colombia looking to migrate to improve her chances of employment. Given this scenario, participants were asked to navigate towards the ‘People like me’ page and to complete the following tasks: 1) identifying the main national and international destinations for people similar to their assumed character 2) assessing the job prospects, income levels, and cost of living in the most frequent destinations where people like them 3) assessing whether or not they felt that the information presented reduced uncertainty in making a migration decision. They were further asked what information would significantly aid in their decision-making process.

Following the tasks, participants filled out a survey evaluation based on the rubric outlined in [Tec18]. The final question examines the tool’s ability to adequately convey information to users and their proficiency in using the tool by asking them what the largest flow of migration was out of Colombia between 2015 and 2020. A 10-minute unstructured feedback conversation was then completed.

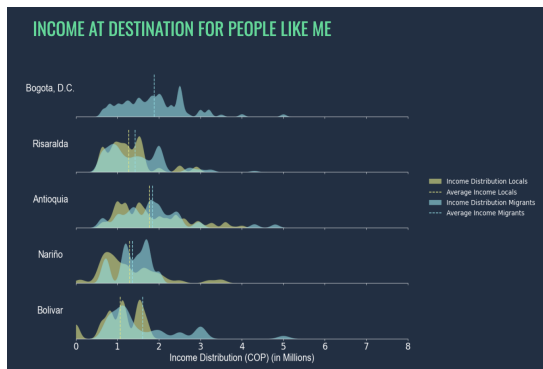
Feedback sessions and surveys were conducted in participants’ native language of Spanish and all participants were graduate students in Colombia. See SI for details on sessions and questions asked.

5.2. Evaluation Results: Strengths

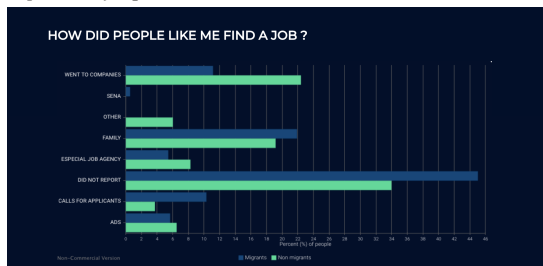
The ability to filter, select and get more information upon hovering over plots supported participants’ insights. On average, this feature was rated 8.5 out of 10. Moreover, one user pointed out that they



(a) Heatmap showing the employment outcomes of migrants that match the user input in the top 5 most frequent destinations. For reference, we included a clickable dictionary with dropdown definitions of employment, unemployed and inactive individuals.



(b) This figure shows the income distribution plots for migrants and locals in the top 5 most frequent destinations.



(c) This bar chart shows the strategies that migrants and locals used to get their jobs in the top 5 most frequent destinations for migrants that match the user input.

would like more fine-grained filters to obtain more detailed information like differentiating people like you by type of job or distinguishing those with Masters from PhDs instead of aggregating them in one category. Data Freshness and data quality were highly rated features in the app, averaging 9 out of 10. The app's usability was rated 8.88, and the performance was rated at 10 out of 10.

During the unstructured feedback conversations, 75% of participants explicitly said they considered the app to be useful, and one remarked on the pleasant aesthetics of the plots. No comments about the data quality, usability or app performance came up during the unstructured feedback conversations.

5.3. Evaluation Results: Weaknesses and Future Work

Data Extension. A significant point where most users agreed on was the need to provide more indicators about the different jobs taken

by migrants at their destinations, crime rates, weather and cost of living in destination communities. One participant commented on the need to include testimonies to contextualize the lived experience of migrants. As a result of the user feedback, visualizations showing the cost of living per destination, and the tone of news about migrants were included on Voyage Viewer. Feedback from the accessibility experts led to changes in font styles and sizes, a revised color palette and appropriate use of CSS labels to make the dashboard usable with text-to-speech tools.

In the future we would like to incorporate other social indicators like diversity of migrant communities living in destination communities. Other avenues for future work include extending the app to more languages, and

5.4. Online Pilot Evaluation

In addition to qualitatively evaluating the use of the tool with a focus group, a small pilot survey was performed on the online participant recruitment service Prolific. The standard sample included ($n = 200$) US participants. Participants were shown information about hypothetical cities names in the same format as the Voyage Viewer platform (income distribution, where people typically migrate too, etc.). Hypothetical city names were used to avoid pre-existing biases towards real life cities. Control participants were told the visualizations represented general trends for the US population, while the treatment group were told the information visualized was about people like them, determined by previous demographics at the start of the survey. At the small scale of the study, no significant differences were found between the control and treatment groups with respect to subjective evaluation metrics like confidence in the choices made. The treatment group had a better performance in information recall, when asked about income data at the end of the study. This effect was not significant given the small sample size. This pilot study highlights the challenge in quantitatively assessing the effect of data on migration decisions. This area is ripe for future work to understand how best to measure the effects that complex personalized data has on decision making.

6. Conclusion

This work has produced a new open-access multi-stream data dashboard aimed at empowering migrants to make decisions about their journeys and relocation. These aggregated and anonymous insights bring together siloed public data to help navigate uncertainty in the global migration context. The tool was evaluated by users and accessibility experts and found to be useful to potential migrants, although the majority also expressed the need for richer data about detailed aspects of migrant outcomes. This tool is a significant move towards creating a reliable information source that can help bring together data to support one of the most important features of our globalized modern society.

7. Acknowledgments

The authors want to thank Vincent Le Put for his work and advice in the platform's development. Thanks to Bernardo Garcia Bueno Bulle, Katherine Pan, and Professors Selim Balcisoy and Cesar Uribe for insightful comments.

References

- [AAB*13] ANDRIENKO G. L., ANDRIENKO N. V., BAK P., KEIM D. A., WROBEL S.: Visual analytics of movement. In *Springer Berlin Heidelberg* (2013). 2
- [AAFW16] ANDRIENKO G., ANDRIENKO N., FUCHS G., WOOD J.: Revealing patterns and trends of mass mobility through spatial and temporal abstraction of origin-destination movement data. *IEEE transactions on visualization and computer graphics* 23, 9 (2016), 2120–2136. 2
- [Bar10] BARNUM C. M.: *Usability testing essentials*. Elsevier, 2010. 2
- [BAS*20] BAHRAMI M., ANDERSON J., SOUTH T., HARDJONO T., PENTLAND A.: The atlas of opportunity, 2020. URL: <https://opportunity.mit.edu/>. 2
- [BCG*17] BELTRAN R., CORNEHLS S., GILARDI F., GUT R., SCHEIWILLER P., STEPANSKI L.: Mapping global refugee movements, 2017. URL: <https://refugeemovements.com/>. 2
- [BFH*18] BANSAK K., FERWERDA J., HAINMUELLER J., DILLON A., HANGARTNER D., LAWRENCE D., WEINSTEIN J.: Improving refugee integration through data-driven algorithmic assignment. *Science* 359, 6373 (2018), 325–329. 1
- [BHCM19] BERRYHILL J., HEANG K. K., CLOGHER R., MCBRIDE K.: Hello, world: Artificial intelligence and its use in the public sector. *OECD Working Papers on Public Governance No. 36* (2019). 1
- [Bur] BUREAU U. C.: American community survey public use microdata samples. URL: <https://data.census.gov/cedsci/>. 2
- [CBG21] CILALI B., BARKER K., GONZÁLEZ A. D.: A location optimization approach to refugee resettlement decision-making. *Sustainable Cities and Society* 74 (2021), 103153. 1
- [CCCS19] CURRY T., CROITORU A., CROOKS A., STEFANIDIS A.: Exodus 2.0: crowdsourcing geographical and social trails of mass migration. *Journal of Geographical Systems* 21, 1 (2019), 161–187. 2
- [CLS] CARPANELLI M., LI L., SCHWARZ M.: Syria visualised project. URL: <https://www.syria-visualized.com/>. 2
- [Com17] COMTRADE U.: International trade in goods and services based on UN Comtrade data. Retrieved from *International trade in goods and services based on UN Comtrade data*: <https://comtrade.un.org/labs/dit-tradevis> (2017). 2
- [DB15] DIJSTELBLOEM H., BROEDERS D.: Border surveillance, mobility management and the shaping of non-publics in europe. *European Journal of Social Theory* 18, 1 (2015), 21–38. 1
- [dED] DE ESTADISTICA (DANE) D. N.: Gran encuesta integrada de hogares microdatos colombia. URL: https://microdatos.dane.gov.co/index.php/catalog/MICRODATOS/about_collection/23. 2
- [DED19] DUAN Y., EDWARDS J. S., DWIVEDI Y. K.: Artificial intelligence for decision making in the era of big data—evolution, challenges and research agenda. *International Journal of Information Management* 48 (2019), 63–71. 1
- [fRU16] FOR REFUGEES (UNHCR) U. N. H. C.: Unhcr statistics: The world in numbers, 2016. 2
- [ftCoHA] FOR THE COORDINATION OF HUMANITARIAN AFFAIRS U. N. O.: The centre for humanitarian data. 2
- [GDE] GDELTPROJECT: Gdelt 2.0 event database. URL: <https://www.gdeltproject.org/>. 2
- [Guo07] GUO D.: Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographical Information Science* 21, 8 (2007), 859–877. 2
- [HP19] HARDJONO T., PENTLAND A.: Data cooperatives: Towards a foundation for decentralized personal data management, 2019. [arXiv: 1905.08819](https://arxiv.org/abs/1905.08819). 1
- [Ije14] IJEOMA E.: The refugee project, 2014. URL: <https://www.therefugeeproject.org>. 2
- [KCB21] KUTTLER E., CILALI B., BARKER K.: Destination selection in environmental migration with topsis. In *2021 Systems and Information Engineering Design Symposium (SIEDS)* (2021), IEEE, pp. 1–6. 1
- [KKHM16] KAY M., KOLA T., HULLMAN J. R., MUNSON S. A.: When (ish) is my bus? user-centered visualizations of uncertainty in everyday, mobile predictive systems. In *Proceedings of the 2016 chi conference on human factors in computing systems* (2016), pp. 5092–5103. 3
- [Las15] LAST T.: Deaths at the borders: Database for the southern eu, 2015. URL: http://www.borderdeaths.org/?page_id=425. 2
- [MA14] MIKSCH S., AIGNER W.: A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics* 38 (2014), 286–290. 2
- [MCDP21] MORO E., CALACCI D., DONG X., PENTLAND A.: Mobility patterns are associated with experienced income segregation in large us cities. *Nature communications* 12, 1 (2021), 1–10. 2
- [MCP19] MORO E., CALACCI D., PENTLAND A.: Altas of inequality, 2019. 2
- [MIC20] MIGRATION, INCLUSION CENTRE M. U.: Australian border deaths database, 2020. URL: https://www.monash.edu/__data/assets/pdf_file/0018/2523141/BOB-Research-Brief-18-border-deaths-annual-report-2020_Final.pdf. 2
- [MK20] MASSO A., KASAPOGLU T.: Understanding power positions in a new digital landscape: perceptions of syrian refugees and data experts on relocation algorithm. *Information, Communication & Society* 23, 8 (2020), 1203–1219. 1
- [NUM] NUMBEO: Numbeo cost of living. URL: <https://www.numbeo.com/cost-of-living/>. 2
- [PLH21] PENTLAND A., LIPTON A., HARDJONO T.: *Building the New Economy: Data as Capital*. MIT Press, 2021. 1
- [SAA*21] SÎRBU A., ANDRIENKO G., ANDRIENKO N., BOLDRINI C., CONTI M., GIANNOTTI F., GUIDOTTI R., BERTOLI S., KIM J., MUNTEAN C. I., ET AL.: Human migration: the big data perspective. *International Journal of Data Science and Analytics* 11, 4 (2021), 341–360. 2
- [SH11] SIMOES A. J. G., HIDALGO C. A.: The economic complexity observatory: An analytical tool for understanding the dynamics of economic development. In *Workshops at the twenty-fifth AAAI conference on artificial intelligence* (2011). 2
- [SLM21] SCHEIDL A., LEITE R. A., MIKSCH S.: VisMiFlow: Visual Analytics to Support Citizen Migration Understanding Over Time and Space. In *EuroVis 2021 - Short Papers* (2021), Agus M., Garth C., Kerren A., (Eds.), The Eurographics Association. doi:10.2312/evs.20211056. 2
- [Smi13] SMITH V. S.: Data dashboard as evaluation and research communication tool. *New directions for evaluation* 2013, 140 (2013), 21–45. 2
- [SO17] SAARINEN V., OJALA J.: The flow towards europe., 2017. 2
- [Spi07] SPIJKERBOER T.: The human costs of border control. *Eur. J. Migration & L* 9 (2007), 127. 2
- [Tec18] TECHBOOM: A rubric for data visualization. <https://medium.com/@techboomph/a-rubric-for-data-visualization-a053b545be32>, 2018. 3
- [Tob87] TOBLER W. R.: Experiments in migration mapping by computer. *The American Cartographer* 14, 2 (1987), 155–163. 2
- [WDS10] WOOD J., DYKES J., SLINGSBY A.: Visualisation of origins, destinations and flows with od maps. *The Cartographic Journal* 47, 2 (2010), 117–129. 2
- [WP11] WEBER L., PICKERING S.: *Globalization and borders: Death at the global frontier*. Springer, 2011. 2