Local Ground: A Toolkit Supporting Meta-representational Competence in Data Science

Sarah Van Wart, University of California, Berkeley, vanwars@ischool.berkeley.edu
Tapan Parikh, University of California, Berkeley, parikh@ischool.berkeley.edu

Abstract: Local Ground is an online data collection, mapping and visualization platform that allows youth to learn and use data skills in support of local projects. Local Ground is unique in its ability to display and translate between a variety of representations of spatial phenomena, starting with drawings. In our studies, we have found that students have used Local Ground to combine and translate between different representations for sense-making and communication.

Introduction

The term data science has been coined to describe the mindset and skills one must possess to acquire, manipulate and use data to understand and predict real-world phenomena, and to communicate that understanding for decision-making. Data scientists play increasingly important roles in a variety of fields - including public health, climate science, governance, the social sciences and for study of the Internet and telecommunications networks. Researchers have observed that data scientists require a unique combination of skills - including statistical reasoning, information organization, and algorithmic thinking (Kandel, Paepcke, Hellerstein and Heer, 2012).

Local Ground is an online data collection, mapping and visualization platform that provides youth with an opportunity to learn and use data skills in support of local civic engagement and citizen science projects (van Wart, Tsai and Parikh, 2010). Local Ground allows learners to 1) collect locally relevant data, including as hand-drawn map annotations, unstructured audio and images and handwritten tables, 2) enter, tag and georeference this raw data into a usable digital format, 3) explore and visualize this data, including as heatmaps, tag clouds, graphs and charts and 4) create compelling narratives from data and multimedia, to be presented and shared with others. We have applied this process of data-driven inquiry across subjects and age groups, from elementary to high school, while working on projects of local civic significance - what Freire calls “generative themes” (Freire, 1970).

Meta-Representational Competence and Data Science

Local Ground is a type of "auxiliary stimulus" - a cultural form that is wedged between students' naive ways of thinking about spatial phenomena and their spontaneous inclinations to represent this thinking. Local Ground is unique in its ability to display and translate between a variety of distinct representations of this phenomena, starting with and-drawing on existing maps. Drawing has been referred to as “probably the biggest and most obvious pool of competence” in children (diSessa, 2004, p. 309). Researchers have also observed that using drawing, students are able to reinvent a number of canonical representations, including graphs and maps (diSessa, Hammer, Sherin and Kolpakowski; Enyedy, 2005). Enyedy refers to this process as progressive symbolization, progressively refining representations of the world to develop a deeper (and more operational) understanding of a domain (Enyedy, 2005).

Data scientists must also create and translate between representational formats as they manage data across the pipeline - from collection, to entry, to modeling, to programming, and finally to communicating results (Kandel et al., 2012). They must integrate data and visualizations into compelling narratives, in the form of presentations, videos and graphics, that can influence decision-makers, the public and even themselves (Segel & Heer, 2010). Enyedy observed how students in a 2nd and 3rd grade class reinvented topographical lines as a solution to representing height in a 2D map (Enyedy, 2005). He describes “how the invention of representational forms by individuals occur as part of a larger social process of creating cultural conventions” (Enyedy, 2005, p. 427).

Case Studies

Participating in Urban Planning

Y-PLAN is a design studio based at the Center for Cities and Schools at UC Berkeley. Y-PLAN facilitated a partnership between Berkeley students (mentors), the City of Richmond, and an eleventh grade U.S. history class, to support youth involvement in creating a design plan for a local park. The students went through a twelve week design process, using Local Ground to collect data about existing assets, hazards and risks; created models, and used this information to advocate for a future neighborhood design plan.

Monitoring Air Quality

EBAYS (East Bay Academy for Young Scientists) is a summer science program for low-income youth, sponsored by the Lawrence Hall of Sciences. Working with E-BAYS, students gathered, analyzed, visualized, and summarized air quality data using Local Ground and other tools, and presented their data and findings to their parents, community members, the Port of Oakland, the transit commission, and at the American Geophysical Union (AGU) conference.

Ground-Truthing Civic Data

I-SEEED is a community organization supporting youth involvement in economic, environmental and educational design. Working with I-SEEED, a local school district and other community organizations, youth from different neighborhoods in Oakland used Local Ground to ground truth an existing "grocery store" dataset. Local Ground enabled young people to submit and geo-reference photos, audio clips, comments, and ratings of the stores they visited, and communicate to the school district that only about a third of the stores in the "grocery store" data set provided even moderately healthy food options.

Preliminary Findings and Future Work

In all of these projects, we have found that students have leveraged different representations of data for sense-making, analysis and communication. In the case of EBAYS, when students presented their findings to parents and community members, they used a Local Ground map to represent the spatial coverage of their study; a time series plot to represent how air quality changed over the course of a train ride; a comparative chart series to represent the striking difference in Air Quality between the Pittsburg and Embarcadero stations, a photograph of the train station to demonstrate dust accumulation, and finally some summary statistics using tabular formats. In the Y-PLAN, when students presented their park design plan at City Hall, they used Local Ground as a way to show future visions of the neighborhood, juxtaposed with a photograph in which an existing methadone clinic was sited next to a proposed playground. These representations were later combined in a PowerPoint presentation that also included poetry, spoken word, graphic posters, etc. In the case of I-SEEED, youth used a combination of drawings, pictures and tables to collect data about grocery stores. Based on this data, mentors helped students to create a visualization that colored store locations based on their categories. Later, they made a multimedia presentation combining audio and images to communicate their qualitative findings.

We plan to conduct further studies of the use of representations to collect and understand empirical data. Working with students in grades 4-8 at two local schools, we will be observing how students appropriate the affordances of the tool to create new representations of space and spatial phenomena. We are also interested in how students convert these representations into more usable and communicable forms, how and if those forms are appropriated by others, and how that process supports youth's understanding of core mathematical, statistical and computational constructs. We also want to understand the importance of locally collected data on youth motivation and identity.

References

diSessa, A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, 10, 117-160.

diSessa, A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22(3), 293-331.

Enyedy, N. (2005). Inventing mapping: Creating cultural forms to solve collective problems. *Cognition and Instruction*, 23(4), 427-466.

Freire, P. (1970). *Pedagogy of the Oppressed*. New York: Continuum Publishing Group.

Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2012). Enterprise data analysis and visualization: An interview study. *Proc. IEEE Visual Analytics Science & Technology*.

Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. I*EEE Transactions on Visualization and Computer Graphics*, 16(6), 1139-1148.

Van Wart, S., Tsai, K. J., & Parikh, T. (2010). Local Ground: A paper-based toolkit for documenting local geo-spatial knowledge. *Proc. of the First ACM Symposium on Computing for Development*

Acknowledgments

We would like to thank I-SEEED, Lawrence Hall of Science, E-BAYS, Y-Plan, the UC Center for Cities and Schools, all of our other collaborators and especially all of the students that have participated in our work. This material is based in part upon work supported by the National Science Foundation under Grant Number IIS-1319849.