# Sidewalk Gallery: An Interactive, Filterable Image Gallery of Over 500,000 Sidewalk Accessibility Problems

MICHAEL DUAN*, Allen School of Computer Science, University of Washington, micdun@cs.uw.edu

AROOSH KUMAR*, Allen School of Computer Science, University of Washington, arkumar@cs.uw.edu

MICHAEL SAUGSTAD, Allen School of Computer Science, University of Washington, saugstad@cs.uw.edu

AILEEN ZENG, Allen School of Computer Science, University of Washington, aileenz@cs.uw.edu

ILIA SAVIN, Human Centered Design and Engineering, University of Washington, isavin@uw.edu

JON E. FROEHLICH, Allen School of Computer Science, University of Washington, jonf@cs.uw.edu

What do sidewalk accessibility problems look like? How might these problems differ across cities? In this poster paper, we introduce Sidewalk Gallery, an interactive, filterable gallery of over 500,000 crowdsourced sidewalk accessibility images across seven cities in two countries (US and Mexico). Gallery allows users to explore and interactively filter sidewalk images based on five primary accessibility problem types, 35 tag categories, and a 5-point severity scale. When browsing images, users can also provide feedback about data correctness. We envision Gallery as a tool for teaching in urban design and accessibility and as a visualization aid for disability advocacy.

**CCS CONCEPTS** • Human-centered computing → Accessibility → Accessibility systems and tools

**Additional Keywords and Phrases:** Mobility, disability, sidewalks, crowdsourcing, open data

**ACM Reference Format:**

Michael Duan, Aroosh Kumar, Michael Saugstad, Aileen Zeng, Ilia Savin, Jon E. Froehlich, 2021. Sidewalk Gallery: An Interactive, Filterable Image Gallery of Over 500,000 Sidewalk Accessibility Problems. In The 23rd International ACM SIGACCESS Conference on Computers and Accessibility, October 18–22, 2021, Virtual Event. ACM, New York, NY, USA. ACM ISBN 978-1-4503-8306-6/21/10. https://doi.org/10.1145/3441852.3476542

## 1 Introduction

There is growing recognition of the importance in collecting and publishing open datasets in the accessibility community [6,8,9]. In a literature review of 506 accessibility papers published at CHI and ASSETS over the past decade, Mack *et al.* [9] found that only seven (1.4%) claimed a dataset contribution and, consequently, called on the community to address this gap. Kacorri *et al.* [8] note how other fields such as machine learning and public health have pushed towards open data, enabling better study replication, performance benchmarking, and supporting scientific exploration and discovery. They emphasize, however, that open datasets alone are insufficient and that data-centric tools are also required to expose datasets, broaden reach, and lower barriers to use.

Drawing on recent work in creating interactive tools for open datasets to support science and education—such as *VizWiz's DataSet Browser* [2,17], Kacorri *et al.*'s *IncluSet* [7,8], and the *Cornell Lab of Ornithology's* interactive gallery of bird species [1,14]—we introduce *Sidewalk Gallery*, an interactive, filterable gallery of over 500,000 crowdsourced sidewalk accessibility images drawn from seven cities in two countries (US and Mexico). To our knowledge, Gallery is the first and largest interactive gallery of categorized sidewalk accessibility problems in existence. While professional

* Michael and Aroosh contributed equally to this work.

image galleries like Getty Images provide high-resolution stock sidewalk photos, most of these are pristine, do not include accessibility-related categorizations, and are not focused on access infrastructure. Indeed, although a Getty Image search of "sidewalks" returns 129,653 images, there are only 816 results for "broken sidewalks", 27 for "curb ramps", and 6 for "ADA sidewalks". In research, others have collected and curated large-scale urban datasets such as *CityScapes* [3] and *Places* [19]; however, these initiatives are focused on training computer vision algorithms rather than presenting interactive views of inaccessible sidewalks. One recent exception is *Urban Mosaic* [10], which allows users to interactively query 7.7 million street-level images in NYC and explore spatio-temporal data relationships (*e.g.,* census, crime, zoning); they present a brief case study tracking tactile strip use in NYC over time.

In this poster paper, we present Sidewalk Gallery, which allows users to explore and interactively filter sidewalk images based on five primary label types, 35 tag categories, and a 5-point severity scale. When browsing images, users can also provide feedback about data correctness. While an early prototype, we envision Gallery as a tool for teaching in urban design and accessibility and as a visualization aid for disability advocacy in local communities.

## 2 The Sidewalk Gallery Image + Metadata Dataset

Sidewalk Gallery sources its image and metadata from Project Sidewalk (https://projectsidewalk.org), an open-source web tool for labeling and verifying sidewalk accessibility problems using *Google Street View* (GSV) [4,5,12]. Project Sidewalk uses *gamified missions* to train, engage, and sustain users and to divide and allocate work. In *labeling missions,* users are virtually routed through city streets in GSV to find and label sidewalk problems and assess their severity. In *validation missions*, users review and validate previously labeled imagery through *agree, disagree,* and *unsure* judgments.

Project Sidewalk's labeling ontology is derived from accessible sidewalk standards [15,16] and includes five primary label types and 35 tags. The label types are *curb ramps*, *missing curb ramps*, *sidewalk obstacles*, *surface problems*, and *missing sidewalks*. Each label can contain a severity assessment (1-5 scale where 5 is an impassable barrier for a


**Figure 1.** Example Sidewalk Gallery images showing sidewalk accessibility problems categorized by label type and tag. Label types include *obstacles* (blue circular label), *surface problems* (orange), and *curb ramps* (green). For readability, only one tag is shown per image (bottom-left corner) though often multiple were applied. All images were collected and categorized by Project Sidewalk users from the San Pedro, MX deployment except for "fire hydrant" from Seattle, WA. Not all label types and tags are shown—see Table 1.

| Curb Ramps | Missing Curb Ramps | Sidewalk Obstacles | Surface Problems | Missing Sidewalks |
|---|---|---|---|---|
| Narrow, points into traffic, missing friction strip, steep, not enough landing space, not level with street | Alternate wheelchair route present, no alternate wheelchair route, unclear if curb ramp needed | Pole, tree, vegetation, trash/recycling can, parked car, parked bike, sign, garage entrance, stairs, street vendor, height difference, narrow, fire hydrant | Bumpy, cracks, grass, narrow sidewalk, uneven/slanted, very broken, height difference, brick | Ends abruptly, street has a sidewalk, street has no sidewalks, gravel/dirt road, shared pedestrian and car space |

**Table 1.** The five primary label types and their corresponding tags used in Sidewalk Gallery. Multiple tags can be applied per label.

| | Users | Audit Distance | Curb Ramps | Missing Curb Ramps | Sidewalk Obstacles | Surface Problems | Missing Sidewalk | Total Labels | Total Labels Validated | Label Accuracy |
|---|---|---|---|---|---|---|---|---|---|---|
| **Washington DC** | 1,395 | 5,482 km | 150,680 | 19,792 | 22,264 | 8,964 | 45,395 | 247,095 | N/A | N/A |
| **Seattle, WA** | 3,052 | 2,300 | 51,140 | 26,970 | 7,524 | 17,642 | 26,535 | 129,811 | 68,053 | 82.5% |
| **Newberg, OR** | 242 | 229 | 4,283 | 1,939 | 916 | 1,992 | 6,827 | 15,957 | 7,332 | 88.9% |
| **Columbus, OH** | 380 | 318 | 11,552 | 875 | 2,768 | 4,401 | 5,778 | 25,374 | 9,299 | 86.7% |
| **Pittsburgh, PA** | 68 | 169 | 5,315 | 586 | 2,983 | 2,310 | 2,073 | 13,267 | 1,865 | 87.2% |
| **San Pedro, MX** | 1,105 | 1,285 | 4,006 | 20,090 | 48,080 | 20,949 | 7,167 | 100,292 | 24,197 | 84.1% |
| **Azcapotzalco, MX** | 358 | 93 | 1170 | 1,141 | 1,590 | 2,119 | 293 | 6,313 | 3,784 | 89.1% |
| **Totals** | **6,600** | **9,875 km** | **226,976** | **70,252** | **84,535** | **56,258** | **93,775** | **531,796** | **114,530** | 83.9% |

**Table 2.** Sidewalk Gallery's labeled image dataset from Project Sidewalk. To calculate label accuracy, we divide the number of *agree* validation judgments by the sum of agree and disagree judgements. The Washington DC pilot did not include crowdsourced validations and thus, has "N/A" for those cells. Sidewalk Gallery is currently deployed in all cities but DC.

wheelchair user), an optional open-text description, and one or more label-specific tags. For example, *surface problems* can be tagged with eight additional descriptors, including *bumpy*, *cracks*, and *narrow* (Table 1). All labels include additional metadata such as the image date, the labeling timestamp, validation information, and geo-location (lat/long). Figure 1 shows a subset of sidewalk labels and tags available in Sidewalk Gallery.

Since its 2018 pilot deployment in Washington DC, Project Sidewalk has expanded into six additional cities, including two in Mexico. In total, 6,600 users have labeled 540,000 sidewalk accessibility problems across 10,000 km of city streets and provided 171,000 validations (Table 2). Sidewalk Gallery provides an interactive, filterable image gallery of this data.

# 3 Sidewalk Gallery

Sidewalk Gallery supports rapid visual exploration of the Project Sidewalk dataset, showing live GSV images from the Google Maps API with Project Sidewalk metadata. Drawing on Shneiderman's visual information seeking mantra of "*overview first, zoom and filter, then details on demand*" [13], the primary UI presents a gallery view of labeled sidewalk accessibility images (cards), which can be interactively filtered based on *label type*, *tag,* and *severity level* using the sidebar (Figure 2). Currently, we show nine images per page—users can navigate to previous and additional results using the *navigation* controls at the page's bottom. All interfaces support both English and Spanish—the Spanish translations were completed by our Mexico-based NGO partner, *Liga Peatonal.* Below, we enumerate primary design principles and tasks.

**Serendipitous discovery.** On initial load, Gallery provides a random assortment of images to encourage impromptu discovery. With filters applied, Gallery shows all matching results in a randomly order paginated set.

**Performant filtering.** User interactions with the sidebar are translated into a combination of frontend and backend filtering and database query operations. Results are nearly instantaneous allowing for interactive exploration and refined filtering. Currently, multiple tags and severity levels can be selected per label type, which result in OR operations.

**View and validate.** As Sidewalk Gallery presents crowdsourced data, it is imperfect and subjective. Users can rapidly view and validate data in both the minimized and expanded cards. Validations are uploaded back to Project Sidewalk.

**Detail on demand.** To provide "*details on demand*", images can be clicked to show an expanded view with additional metadata and an interactive GSV pane where users can explore problems by *panning*, *zooming*, and *moving*.

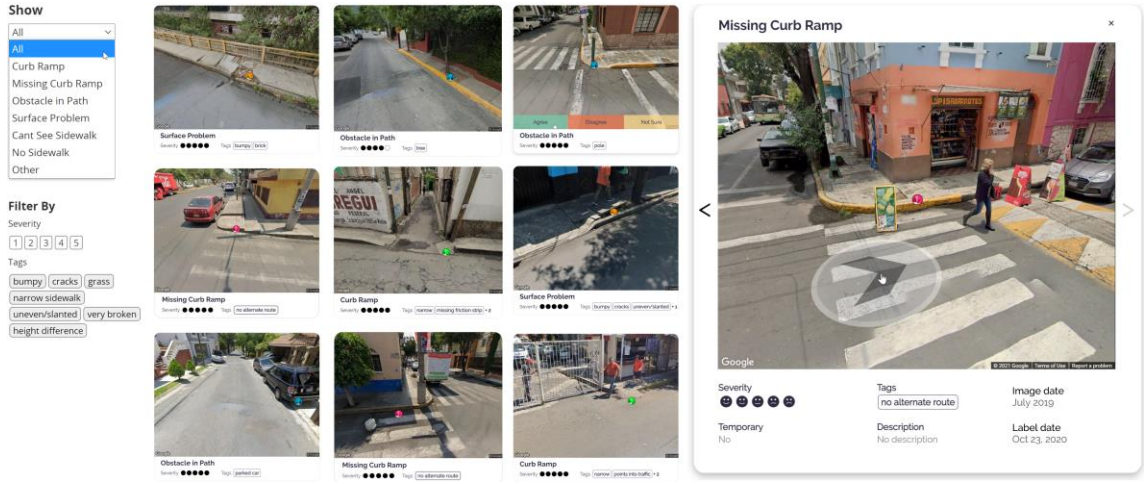


**Figure 2.** The Sidewalk Gallery interface showing the filter bar (left), the main gallery (middle), and the expanded card UI with an interactive GSV pane and more detailed metadata (right). In the gallery view, users can mouse over images to vote on data correctness. For example, in the gallery grid above, the user is mousing over the top-right card and shown an overlay of Agree (green), Disagree (red), and Not Sure (yellow) buttons. In this case, the label is correct so the user can select "Agree".

Sidewalk Gallery's frontend is implemented in HTML5, CSS3, and JavaScript with the i18next framework for multi-language support. The backend is implemented in Scala with a PostgreSQL database. Visit https://sidewalkgallery.io.

# 4 Limitations and Future Work

While Sidewalk Gallery is currently deployed online, it is still an early research prototype. Significant work remains.

**Cross-city comparisons and filtering.** What do the most severe sidewalk accessibility problems look like in city X *vs.* city Y? How do curb ramp designs differ across cities and regions? Future work should explore solutions that allow users to filter and compare across cities.

**What are the geographic patterns of sidewalk inaccessibility?** Sidewalk Gallery does not currently surface the geographic location of labeled images. We plan to address this limitation in two ways: first, by adding a small top-down map on the filtering sidebar, which will immediately show the location of all labels matching the selected filters. Second, by adding a larger version of this map in the expanded view, which will also highlight the location of the selected image.

**Showing and filtering on "correctness."** Sidewalk Gallery relies on crowdsourced data. While roughly 40% of all Project Sidewalk labels have user-contributed validations, this information is not yet surfaced in Gallery. To enable showing the highest-quality data, we plan to allow filtering based on validations and to show validation data in the UI. We plan to also experiment with automated quality inference approaches, which would allow the user to see and/or sort by our trained machine learning predictions of correctness [18].

**Crowd editable metadata.** While Gallery users can currently validate cards, the metadata itself is read-only. Future work should explore methods to make the metadata (*e.g.,* severity, tags) crowd editable.

**Complementary analytic visualizations.** What are the distributions and co-occurrence patterns of label types, severity ratings, and tags? How do they correlate to socio-economic data (*e.g.,* race, real estate pricing, census tract data)? To support addressing each of these questions, we would like to add in complementary visualizations. The filter bar, for example, could show histograms of severity ratings based on label type.

**Image ownership and licensing.** Unlike the VizWiz DataSet Browser [2,17], which shows user-contributed photographs, Sidewalk Gallery serves labeled GSV imagery using the Google Maps API. These images are owned by Google and occasionally "expire" over time. While our geo-located metadata would persist, the images may not.

**User study.** While we have received informal feedback about Sidewalk Gallery from our partners, we plan to conduct a multi-stakeholder user study [11] of disability advocates, people with mobility disabilities, policy makers, and planners exploring how Sidewalk Gallery can be used for sidewalk accessibility analysis, education, and training.

In conclusion, this poster paper contributes a new interactive tool, called Sidewalk Gallery, for exploring and filtering sidewalk accessibility images. Our work has implications for the growing open data movement in accessibility research and should benefit urban planners and accessibility researchers. We also believe Gallery could be a beneficial educational tool, just as the Cornell Lab's Ornithology gallery has been successfully used in zoology education [14].

## ACKNOWLEDGMENTS

This work was partially funded by the USDOT Pacific Northwest University Transportation Center (PacTrans) and the UW Center for Research and Education on Accessible Technology and Experiences (CREATE). We thank Liga Peatonal for their Spanish translations and support, the Project Sidewalk team, and our partners.

## REFERENCES

[1] Ileana Betancourt and Colleen M McLinn. 2012. Teaching with the macaulay library: an online archive of animal behavior recordings. *Journal of microbiology & biology education* 13, 1: 86–88. https://doi.org/10.1128/jmbe.v13i1.382

[2] Nilavra Bhattacharya and Danna Gurari. 2019. VizWiz Dataset Browser: A tool for visualizing machine learning datasets. *arXiv preprint arXiv:1912.09336*. Retrieved from https://arxiv.org/pdf/1912.09336.pdf

[3] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[4] Jon E Froehlich, Mikey Saugstad, Edgar Martínez, and Rebeca de Buen Kalman. 2020. Sidewalk Accessibility in the US and Mexico: Policies, Tools, and A Preliminary Case Study. In *CSCW2020 Workshop on Civic Technologies: Research, Practice, and Open Challenges*.

[5] Jon E Froehlich, Mikey Saugstad, Manaswi Saha, and Matthew Johnson. 2020. Towards Mapping and Assessing Sidewalk Accessibility Across Socio-cultural and Geographic Contexts. In *AVI2020 Workshop on Data4Good: Designing for Diversity and Development*.

[6] Danna Gurari, Qing Li, Abigale J Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P Bigham. 2018. VizWiz Grand Challenge: Answering Visual Questions From Blind People. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[7] IncluSet. 2021. IncluSet Data Browser. Retrieved June 11, 2021 from https://incluset.com/datasets

[8] Hernisa Kacorri, Utkarsh Dwivedi, Sravya Amancherla, Mayanka Jha, and Riya Chanduka. 2020. IncluSet: A Data Surfacing Repository for Accessibility Datasets. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '20). https://doi.org/10.1145/3373625.3418026

[9] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by "Accessibility Research"? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA. Retrieved from https://doi.org/10.1145/3411764.3445412

[10] Fabio Miranda, Maryam Hosseini, Marcos Lage, Harish Doraiswamy, Graham Dove, and Cláudio T Silva. 2020. Urban Mosaic: Visual Exploration of Streetscapes Using Large-Scale Image Data. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), 1–15. https://doi.org/10.1145/3313831.3376399

[11] Manaswi Saha, Devanshi Chauhan, Siddhant Patil, Rachel Kangas, Jeffrey Heer, and Jon E. Froehlich. 2020. Urban Accessibility as a Socio-Political Problem: A Multi-Stakeholder Analysis. In *Proceedings of the ACM on Human Computer Interaction (PACM HCI); To Be Presented at CSCW2020*.

[12] Manaswi Saha, Michael Saugstad, Hanuma Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, and Jon E. Froehlich. 2019. Project Sidewalk: A Web-based Crowdsourcing Tool for Collecting Sidewalk Accessibility Data at Scale. In *Proceedings of CHI 2019*.

[13] Ben Shneiderman. 2003. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *The Craft of Information Visualization*, BENJAMIN B BEDERSON and BEN SHNEIDERMAN (eds.). Morgan Kaufmann, San Francisco, 364–371. https://doi.org/https://doi.org/10.1016/B978-155860915-0/50046-9

[14] The Cornell Lab of Ornithology. 2021. Macaulay Library. Retrieved June 11, 2021 from https://search.macaulaylibrary.org/catalog?mediaType=p

[15] U.S. Access Board. 2013. *Public Rights-of-Way Accessibility Guidelines*. Retrieved from https://www.access-board.gov/prowag/

[16] US Department of Justice. 2010. *2010 ADA Standards for Accessible Design*. Retrieved from https://www.ada.gov/regs2010/2010ADAStandards/2010ADAstandards.htm

[17] VizWiz. 2021. VizWiz Dataset Browser. Retrieved June 10, 2021 from https://ivc.ischool.utexas.edu/VizWiz_visualization/view_dataset.php

[18] Galen Weld, Esther Jang, Anthony Li, Aileen Zeng, Kurtis Heimerl, and Jon E Froehlich. 2019. Deep Learning for Automatically Detecting Sidewalk Accessibility Problems Using Streetscape Imagery. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '19), 196–209. https://doi.org/10.1145/3308561.3353798

[19] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2018. Places: A 10 Million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40, 6: 1452–1464. https://doi.org/10.1109/TPAMI.2017.2723009