WA State Unified Accessible Pedestrian Data Collection Advisory Meeting

📖 Tue, Oct 01, 2024 3:02PM 🕐 1:25:05

SUMMARY KEYWORDS

roadside sidewalks data, project introductions, data collection pipeline, quality control, annual report, priority regions, analytic paths, community deep dives, sidewalk accessibility, data visualization, stakeholder feedback, sidewalk condition data, pedestrian safety, data stewardship, walk shed analysis

SPEAKERS

Anat Caspi, thomas craig

A **Anat Caspi** 00:53

Go ahead and share my screen. I All right, so in a minute, we're going to start the sixth advisory group meeting for the Washington State proviso to collect roadside sidewalks data in Washington State. I am excited to see everybody who's here you Yeah. All right, it is four past the hour, and thanks again for joining. In a minute, we'll have an opportunity to kind of popcorn around the room and see who's here. As an overview, I will do some team introductions today, and hope to gage who is here, we have previously proposed work streams for the proviso work. I'll kind of summarize the current state of the work and where we are and how you may be able to gain access to our ongoing progress, we'll talk about the specific pipeline for data collection as well as the quality control that we've devised, and additional metrics that we'll be putting through both towards the end of the biennium as well as the hopefully in the next biennium, and talk about our next meeting, which will focus on the Annual Report specific priority regions of interest for sidewalk deep dives, and also clarifying the workflows or analytic paths that we want to work on with you stakeholders, to better understand which workflows would best integrate with the type of work that you're doing, as well as what would contribute most to the work that you're doing. So that is the overview of what we'll talk about today, first project team and introductions. So there are several members of our team here today. I function as the project manager and a scientific lead, Ricky Jiang, who recently is still on the project, but out works as the computer vision pipeline development lead, and Olivia Quesada is the partnerships and communications lead. So both have been working on the U DUB side, and Vicki recently moved over to the Gaussian side. On the Gaussian side, the team is led by Suresh devapoli and Pradeep paida. Oh, I'm so sorry, there's a missing P there is our interim Project Manager for this project. In addition, we have Sam Jan working on some of the analytic and visualization tools, and he's been functioning as the development lead for those tools. And Bill Howe has been our scientific advisor. So the group, the team, breakdown as far as responsibilities are concerned, the execution team is focused on pipeline hardening, making sure the imagery is ingested appropriately and vetted. The manual vetting, which is a huge part of the project, ensuring that the predictions through the model are manually vetted and assessed. Reporting tools, as well as the community vetting application that we've been using. And on our side, the OU dub side, we're developing further, the model for the prediction, the post processing for those predictions, doing the quality assessments, the post manual vetting,

ensuring that the data is published, and producing these visualization tools and analytic tools that we will talk about at the second part of the meeting today, also commenting and quality control tools. So how you as our stakeholders and experts are able to comment back on the data to make sure that we are following rigorous protocols, and also a whole other part of this project is doing regional deep dives with communities to engage people in The community in both vetting the data and also providing us with additional inputs on what kinds of attributes must be collected in their own locales because of specific barriers that they experience in the public right of way I wanted We have, I think, enough time to go through and do short introductions. So if you don't mind, maybe I'll call on people, but we if you could just say your name, your organization and hope or goal for this project, that would be very helpful for us. So I will start out with the and with IDA.

07:15

I'm here.

A Anat Caspi 07:18

Hello, if you could just briefly give your organization department and hope goals for this project,

07:25

skull bank, I'm with Washington State Department of Transportation. My group are the asset stewards for State Highway, sidewalk programs and grass walks. And my hope a goal for the project is that will have greater coverage of the state of Washington for these kind of features,

A Anat Caspi 07:48

hoping for complete coverage. Yes, thank you. John Deskins,

T thomas craig 07:56

Hi, I'm John Deskins.

07:58

I'm the city traffic engineer for Richland, Washington,

08:01

and I've been invited, so I don't really know a lot about the project yet. Certainly, we are trying to

08:09

track and catalog all of our ramps and sidewalks as well, so hopefully we can contribute in some way. Great

Anat Caspi 08:19

to see you. Thanks for joining. Justin

08:24

Denno. Justin Denno, cm, pronouns, found transit. I work within the research and innovation group. I think my long term hope for this project and any sort of project like this, is that the data that we collect that we're able to communicate to pedestrians and passengers, you know, as closely to the reality on the ground as possible. I think that's the hope that we can feed these things into our tools and give people picture on the sidewalks and pedestrians.

09:04

Thank you.

Anat Caspi 09:06

Thank you. Justin Benjamin klosky,

09:11

hi, I'm an associate transportation planner with Spokane Regional Transportation Council. I guess I'm just really here to learn more about this project. If I had one hope that I would share is that just as this project gets developed and in the years coming, it just continues to be maintained, and having an up to date sidewalk inventory over the coming years would be great. Thanks.

Anat Caspi 09:31

Thank you. I'm going to skip over all our internal team members, so that included skipping Sam Reina and Jeff Mach, Thomas, Craig,

thomas craig 09:45

Hi, I'm Thomas. Craig, I use he, him pronouns. I@wash.in'm the public transportation division.

And I guess hope or goal for this project is statewide consistency.

Anat Caspi 10:01

That's a great goal. Grace,

10:06

hi everyone. Grace, young, she her pronouns. I'm also with washdot in the Active Transportation division, and my hope is a really nice statewide, accurate data set.

Anat Caspi 10:18

Thank you. Brian Lee,

10:23

hey everybody. Brian Lee, he him. Pronouns. I work for the Puget Sound Regional Council, the MPO for the four County Central Puget Sound region. I'm a program manager in our data department, and I think there are lots of things that I can envision using this data for couple of hopes we do like a piecemeal inventory right now, so we're definitely hoping for a much more extensive, perhaps complete, inventory of these facilities in our region, and that could support a whole variety of work programs, from more modeling to measures of accessibility and the other meaning of accessibility, access to destinations, but also kind of the more universal meaning of accessibility as well. Thanks.

Anat Caspi 11:10

Thank you. And we have been looking at some of the data that you said, as you said, collected piecemeal, helpful. So thank you for that. Michael redlinger,

11:25

hi folks. Michael redlinger, with Spokane, regional transportation Council colleague of Ben here, I have been attending these meetings to stay up to date on this project. We would love to, at some point, be able to use this data to help us with our region wide sidewalk data, and maybe come arrive at an accurate sidewalk inventory and see how this can be applied to the regional active transportation networks. And you know that it opens up a lot of possibilities for us as far as pedestrian planning and active transportation planning in general go.

Anat Caspi 11:59

Thank you. Scotty that's

12:01
all we

12:04
have. Hello. I'm Scotty Alton from the City of Spokane. I just said this many meeting invite, so I'm trying to learn about it. I do GIS for the City of Spokane and created the sidewalk network that is available.

Anat Caspi 12:16
Awesome. Thank you, Kathy. I

12:23
Hi, Kathy Fitzpatrick, I'm with the mid Columbia Economic Development District. I serve a by state region, and my hope for this project is is to is to get people on transit safely and comfortably.

Anat Caspi 12:47
I'm so glad to have transit interests here. Thank you. That's to Justin as well. Kevin

12:58
Sure. My name is Kevin Bacon, so principal engineer the City of Spokane Scotty just introduced themselves. We worked together. Just recently, became aware that this work was going on, and got forwarded the invite. So participating just to understand better what is going on and how it might apply to our needs. Lovely.

Anat Caspi 13:17
Thank you. Julie Jackson, you

13:23
Oh, I am. I am not

13:24

muted. And can you hear me? Yeah, I am. Julie Jackson with Washington State Department of Transportation. Just here is a fly on the wall to hear a little bit about where this project is at and how we might integrate it with existing wash dot data workflows,

Anat Caspi 13:44

awesome, fantastic. So that was everyone who's part of our team. Do you want to quickly just show yourselves and say hi. Sam has already been on camera, great. So everyone's working super hard on this project, and so we're super excited to tell you what's been going on. Are you able to see my slide? Yes, I hope sorry I don't see thumbs up or anything, so say something fabulous. Okay, so for those of you are new, and some have actually attended each and every one of our five previous meetings. So thank you for your consistency. The project objective is to create a statewide, consistent, connected graph that is routable and openly shared with sidewalk data, the motivation is to further the needs and prioritization of Washington State transportation plan, including supporting activities in active transportation, pedestrian safety, vision, zero, complete streets, sustainable, resilient, ecologically sound communities, and correcting the harms in traditionally underserved communities, including access and reach. And the key deliverables is a Washington State pedestrian dataset in OSW, which stands for Open sidewalks. Data schema, version 0.2, our project scope are areas that coincide with 80% of the state's population. In addition to that, the densest population centers and transit facilities in the missing counties that aren't represented in that coverage of 80% of the state's population, and so overall, 6400 square kilometers are to be mapped. Open sidewalks is a data schema we've been working on since 2016 it is, first and foremost, a transportation layer with a pedestrian graph. So the intent is to demonstrate connectivity of paths and the entities most most vividly represented there for pedestrians are sidewalks, the links between the sidewalks and the crossings and crossings there are now other entities represented in the data, including roadways and bicycle paths. But we're focusing here on the collection of sidewalks, crossings and the connectivity between them. Aside from defining a baseline data schema, we have downstream use cases that use the schema with extensions. So what that means is that we have sort of a level one kind of collection for this base graph, but additional attributes may be used for different use cases. Because, as we said previously, the motivations are many to collect this data. We have interest in Active Transportation and Safety, vision, zero, et cetera. And so different stakeholders might have different types of extensions in mind when looking at that data collection. So the proposed work streams for this work are many, but we summarize them here as first reviewing what jurisdictional data existed so we are aware of the PSRC data set as well as the Spokane data set, and some of you talked about that. We also looked at other data sets that were available. We assessed the state of practice, but also provided our own pipeline for data predictions so that we can do so consistently and using the most up to date satellite imagery that we have access to, which is something I will speak to in just a minute. But the three objectives that we're currently concurrently working on are, one is have local community focus, so we're working with different localities to understand the specific barriers and potential extensions that they might want to work on, collecting to add on to the data so they they can perform certain workflows, like a planning workflow, and we'll talk about that in a minute. Sidewalk collection, the schema and the full collection, including the QC pipeline, is an additional effort that we're putting forward. It includes a large scale effort on doing model prediction and retraining, but also, as I mentioned before, doing a whole lot of manual based assessment of what the model is doing and correcting what it's doing. So there's a lot of error correction there. And thirdly, providing the sidewalk accessibility demonstration applications,

which means showing how that data can be consumed downstream for specific use cases. We'll start some of those demos today. And so we're obviously not quite there at the sidewalk Data Summary and outcomes. But the end of this project for Well, I'll say this. We had originally intended to work on this project for four years the biennium obviously approved us for two and so we're up at the end of this first biennium, at the end of June. We previously detailed the work streams with a lot of additional detail, and I will be sharing these slides afterwards, but I think focusing on these three items for today will give us kind of the level of detail and scope that's needed. But please, for those who are interested, go ahead and review some of our additional documentation, because there is a lot, and we need your help in providing the annual report at the end of this year, December to the legislator. So the timeline up to now has included a project kickoff with the group kickoff, where we did a needs assessment, and many of you who are here today also contributed both the types of workflows you intend to work on this data, but also some of the concerns and barriers that have traditionally impeded the collection of this data. And so we were well advised by your input. We did a lot of data review and best practices work, and continued to work on hardening the computer vision pipeline obtaining the imagery, which was a bigger hurdle than we had imagined, and now the data that is being used for the data collection is the Washington State Code. Collective data, that is essentially the hexagon data. Up till now, we've had access to the 2021 data, and very soon we ought to have access to the 2023 satellite imagery data. And so that is something to keep in mind, because while that provides the predictions to be aligned with the imagery that everybody who works with the state or as a partner of the state has access to. So that's important, because we can't potentially use some of the maxar or other satellite imagery for alignment. On the other hand, the 2021 state of things may not necessarily align with the 2023 state of things. And also, some of the data quality was not at on par with the data quality around more urban centers, et cetera. So that is one of the barriers that we've or challenges that we've had in collecting this data. The second work group meeting took place in March, and that's where we provided an update, as well as an agreement on the schema and the approach. So that map that I showed earlier in terms of where the population centers that will be collected are, and agreement on basically the baseline schema took place in that meeting. Very close to that meeting was an additional meeting to discuss which aerial imagery we ought to work with. And the agreement was on the satellite imagery by hexagon. And then in May, in sorry, in April, work group, meaning four, took place where we discussed Lane cover and as well as some of your recommendations on how to go about some of the deep dives, the localities and where to look at additional barriers that we ought to collect for some of the extensions of the data set. In the meantime, sort of in the execution land of things, we've been working with the machine learning pipeline, as well as human vetting, and at this point, 30% of our intended areas have been covered. We've also worked on the data visualization tools and the ability to comment on the data, basically opening it up for stakeholders like you, to be able to look at the data as it's being progressively added onto and QC. And so our work group reading today is going to talk about that QC pipeline, the visualization and comment tool, and also a specific deep dive that we are currently working on, which is called the health through housing project. And I'll explain more about that in a minute when we cover that. So what is the current state of sidewalk data? I have a QR code here where you can yourself go into that report, but I've also hoping to go into it, because when you open the report, you'll be able to see a county by county analytic of where we are with the data collection. I should say that this is not the complete so when we say ready to release, that's actually from our Gaussian team to the U DUB team, so there are two additional QC steps after this release. So even if it's green, it's not in its final state of publication. So what is in red has not yet been scheduled for the machine learning pipeline. What's in purple, which is actually not very much of it essentially, has been completed, gone through the satellite imagery computer vision pipeline. And predictions are there, but they're currently set up for validation by the manual mappers. And

you'll see manual mappers are currently working on what is in blue. And you can see, for example, if you zoom in, so some areas right here are in blue, and then those areas that are green have gone through both the computer vision predictions, as well as the human mapper vetting and analysis. And so what will happen after this is another quality control and analytics that we do, which is essentially spot checking through walk sheds and trying to better understand whether the connectivity is true to what's on the ground. So feel free to visit this page as well as ask any questions. Oh, I should probably talk a little bit about the numbers. So the percentage is represented as the percentage of the total area we intend to map overall. So that's the coverage for 80% of the state's population. And in addition to that, we will do a different type of analytics for those areas that really didn't have coverage, but are transit hubs and stations and population centers in the counties that may not be represented for 80% of the state's population. So overall, we should reach 86% coverage for the state's population by the end of this biennium. If we're looking at validation in that what's been validated, so it's about 30% of that. We do have two additional cities, those are Seattle and Bellevue that have already been manually vetted throughout, and so they are actually not represented here. So data completion for manual funding would actually be pushed to around 34% if we were to add in those additional areas. And you can examine the sidewalk counts that are represented here, curb counts, crossings, etc. Any questions on this before I move on. I see a chat, maybe I should check it out. Oh, thank you. Olivia, appreciate it. Any questions?

27:54

Hi, Annette. This is Michael. I just wanted to see if you could go into any detail for those green areas on the map, what the timeline for those two QA QC steps might look like, and what some of the dependencies are there?

28:09

Great question. So I will go into is a specific QC that will be implemented soon. But I wanted to mention that you can start viewing the data even immediately, even if it's not at its final state. And we would love to get your comments on it, too. The comment tool, well, the comment function in the tool is not totally there yet, but it's coming. But you can right now go to viewer.sidewalks.washington.edu,

Anat Caspi 28:42

and I will flip to that. And sorry that I'm showing the slides in this janky way. I just didn't want to have to keep flipping windows. So somehow that didn't work out for us. Okay, I'm Are you viewing the TD I viewer? Are you able to see this?

29:13

Yes, perfect.

Anat Caspi 29:14

Okay, so let's go to Spokane, and so you'll see the data versioning by going to any specific one of the cities. We can focus on Spokane, going right in there. Let's say we're interested in this area around Drumheller springs park, you can identify whether the assessment identified that there is or is not a sidewalk. Here, you can click on that sidewalk and it will tell you what the proposed surface there might be. We don't have the inclines calculated for them quite yet. In addition, the crossings will let you know if we assess that there are curb ramps here. Curb ramps means that it ought to be curb ramped on both sides, but we have the data separately for each for each of the potential curves, it will also indicate whether the data had that's been collected, indicates that there is a marking there or not, and the surface would be annotated as well. You can look at the satellite imagery from which this data was gleaned. So again, if we're talking about just the computer vision pipeline, and what the humans who are vetting are able to see, this is the data and imagery that they have access to. And so in many cases at this point, we actually, in all cases, at this point, it's all 2021, data. So if there are areas that have been updated since that time, it will not be up to date. But we are creating the tools to speak to what someone was talking about maintenance. I think it was Kevin about sustainability and maintenance. We're absolutely cognizant of the fact that things change on the ground all the time, and so there is this tooling ecosystem that we're providing to allow stakeholders on the ground, or those who are aware of it, to update the data as we go and as we move into more the of the data stewardship, rather than the collection part of this project. So let's see. And I thank Sam for enabling this tool as quickly as he did. Now, a lot of the input for this tool came from Thomas Craig and Grace young and Mary DeBose, who've been advising us throughout the project. That's not to say that you are not necessarily able to provide us with input. So if you do have comments to share, would love to get them in the chat or send us an email, because this this tool is really for to to enable everybody to comment on the data and also provide us with guidance, whether they see certain flaws or areas that might be flawed.

thomas craig 32:38

I had my hand raised, so I'll go ahead and kind of, you kind of just answered my question. I was going to say basically, sort of because, you know, yeah, as you just indicated, myself and grace and Mary have gotten the sedis tool, and have had some time to get in and take a first look at data sets. And as we went through, you know, we tried to provide some feedback on the viewing interface, and what would it make, and what would make it an effective interface for us to be able to engage with the data, and then, not so much kind of like specific issues, necessarily, though, pointing to specific geographies as examples, but kind of trying to provide some high level feedback on, oh, as we're looking through like we're we're finding these sorts of scenarios seem to be handled effectively, these sorts of scenarios, like the false negative scenario that we discussed a couple of weeks ago. Like, think we're seeing some of this, and tried to kind of provide sort of high level feedback at this point, as opposed to a laundry list. But the question basically is, like, you know, from this group, are you looking for the laundry list at this point? Are you looking for, you know, the high level feedback, like, what's kind of, what's the most useful way at this point in the process where your team is in developing data, what, what's kind of the right level of feedback

34:09

specificity

Anat Caspi 34:13

from our standpoint? Yeah, like, I

thomas craig 34:17

mean, you know, is this? Is this the right time, if someone you know working at at srtcc, you know now has access to this, should they go through and you know, compare to their local inventory and identify specific sidewalks that are missing or not missing, and like provide line by line feedback, or is it more useful at This stage to get sort of, like, higher level feedback on types of issues, or like, what are the what are the types of comments that are most useful for your team?

Anat Caspi 34:48

Yeah, thank you for that. Okay, that's a great question. So as I noted, we're not quite done with the full QC steps, and part of that has to do with identifying correct connectivity on the ground as identified some false negatives issues. So false negatives means that there is a sidewalk there, or a portion of a sidewalk, but it's not identified in the data set.

35:17

Specifically with crossings, our initial directive to the human mappers was actually not to do every single crossing at intersections, specifically looking for potentially like large Boulevard and any any road that had four lanes in it was indicated as a not great place to cross, because we were coming in with the Access Map mindset, which is like, not to lead people potentially into unsafe situations, but realizing that

Anat Caspi 35:52

the the legal crossing, at any crossing, unless indicated otherwise, would say that, there ought to be a crossing indicated there, even if it's not necessarily safe or marked. And so we did have to go back and rework with our human mapping team to ensure that we're now representing those in the data sets. So for example, you can see right here, this would now be marked for crossing, whereas it was possibly even deleted from the predictions previously, because there was really nowhere to cross to, or something like that. So keep those two things in mind, but if you're seeing some areas of concern, if you're seeing areas that are specifically of local concern for you, and you see mismappings There, that's really good for us to know. We're super interested, and that's I'm about to go into, the local deep dives conversation. We're really interested in working with people who are on the ground and have access to being on the ground so that we can further improve the quality of the data. The problem for us is we can't be on the ground everywhere, and so we've been working with teams who are able to provide that that level of you know, assessment, so that we can work with you on improving the data in specific locations and locales. In coming weeks, certain areas will have gone through the full

like QC pipeline. And then we will let you know what I'm thinking we ought to do with a viewer specifically is to, sorry, not with a viewer, but rather with the wisp project. Sorry, this is I have the zoom, the zoom thing just stuck in front of everything else

37:59

here, right here. So if we go back to the reporting, you'd be able to see further where the additional QC had taken place. So it will be an additional step past the green, so that you then it really will be up for very detailed commentary, if you want to say, oh, you know this, this is not concrete, or the incline here is incorrect, or something like that. Did that answer your question?

thomas craig 38:29

I think so totally. So. Yeah, what I'm hearing is maybe number one request, if someone wants to, you know, look at this data and provide feedback, is just to get in touch with your team and actually reach out, start a conversation, you know, hopefully be happy to partner and like, have an ongoing conversation where, you know, you get some access to local resources and a way to contribute in the future. And then, you know, we're not quite ready for the like, you know, the detailed inventory of this, this, this, but that's coming on the horizon, and you'll let this group know when, when we get there, absolutely

39:09

and in fact, I mean, I don't want to inundate people with too many emails, but I we produced that report on a weekly basis. And so when specific areas complete the full pipeline, we can just send an email out for people to check it out.

Anat Caspi 39:32

So talking about community deep dives like what's possible to do, and this image is from our recent community deep dive in Renton, near by the Sydney Wilson housing facility. And so the project stakeholders here are King County Metro's health through housing initiative. It is connected to the King County connected to transit initiative. But we've also gathered the help of the Empower movement, which is a group of stakeholders who self identify as being disabled and bipoc and are working on transit access and hopelink, community transportation navigators, who are representatives that go into these health or housing facilities and try to engage people in access to transportation and community access in general. So these are the people who are working on the ground with us, and all through the leadership of Olivia, who said hi earlier. So our objectives for like the most recent sprint, and this is all, all of this is as of the 26th of September, we have continuous walking audits around 15 different housing facilities, some of which are all of which are in King County, but are spread throughout. There's different facilities in Kirkland, in Seattle, in Renton, in Tukwila, Burien, Federal Way, et cetera. And so we provide people with a mobile app for auditing the data that we've created. But in addition to audits, they provide additional attributes which are considered an extension that King County, connected to transit has defined as their extension of need in order to be able to

assess people's access and connectivity to transit. And so that kind of collection adds on, in addition to the data that you'll see in the viewer over there, adds additional attributes for the sidewalks and crossings in those regions of interest. And so what we do is we take that data through so in this, in this discussion here, representing which areas we're looking for that deeper dive in. There are 15 different housing facilities, but right here, we're concerned with, I think, 13 of them,

42:30

and adding the data attributes for the schema basically means that we're able to do more complete walk sheds analysis for that, for those areas. So what does that look like? In this particular example, we're focused on a specific location called the Argyle housing facility. It is in downtown Seattle, and what we're seeing is different types of simulated mobility profiles so different types of pedestrians and what their access might be from that specific location and facility. And so in this we are taking into account the sidewalk, the sidewalk connectivity, the crossings, but different mobile mobility profiles will, for example, require curb ramping everywhere. And so that's why you might see that the manual wheelchair user represented by the pink walk shed from Argyle, it has a more limited walk shed then might be the control, which is the presumed like Google Walker, who has no mobility limitations or barriers, and so that you know black represented walk shed is exactly what's considered to be the 15 minute walk shed From Argyle, and might require walking where no sidewalks exist at all. Whereas as we move up through these different mobility profiles, you'll see that more and more constrained mobility profiles will indicate that you have lesser connectivity and access and reach in those areas. So that's not just a representation of what we can do with a baseline data that we're collecting, but also some of the deep dives. And in addition to that, it's also representing what our new walk sheds tool is enabling. So wherever the data is collected, we can also do this walk sheds calculation, and the tool is also going to be up for release in the next quarter. In addition to being able to download the walk sheds based on different mobility profiles, we're also offering the sort of

A   Anat Caspi 45:01

walk shed ISO contours. So the idea being that you'd be able to look at the different walk sheds and identify like the five minute, 10 minute and 15 minute walk sheds, for example, and also be able to see some of the additional information about that walk shed. So what, how many sidewalks are accessible to this particular individual that was represented by the mobility profile? How many crossings, how many unpaved edges are there so that they might be inaccessible to some of the pedestrians? What percent of our marked crossings, things like that, and also having to do with curb ramps, etc. So what you might see here is that, again, you have different type of reach for different types of pedestrians, and it's all you can change the settings in this tool so that you can calculate whatever it is that you want. There's no like presets to constrain the types of walk sheds that are calculated there. So I think that may actually, in addition to being a really cool tool, this too, provides us with a way to QC the data, and potentially for you as well, when we open it up, because it means that you can vet the data on the basis of the routing or walk shed calculation that's there, and be able to kind of assess out whether certain connectivities have been missed, for example, or whether some connectivity has been conferred as there when it is not there. And so we've used that as a very useful QC tool for some of the data. Whereas, like looking at the viewer you want, you have to

assess every kind of connection on its own using the routing tool, you're actually able to see like, whether the whole thing is actually connected or is flawed. Its connectivity is flawed in some way. And I argue sometimes that connectivity is much more important than like, the asset collection itself. Because, you know, assessing reachability is not just about like whether there's a curb ramp there or not, but whether the entire like a specific location is connected enough to its vicinity. I this is kind of a lot, so I want to open it up for others to ask questions. I

47:49

Yeah, okay, and that, yeah, I don't I hope this doesn't take us back, because I've missed at least, I think one meeting before. I'm just kind of curious. Are? Is there a distinction made between sidewalks and infrastructure that's maintained by public agency versus facilities that are publicly accessible, but they're more quasi public spaces, but it's open to pedestrians to use.

Anat Caspi 48:28

I think the latter would be like trails. Is

48:33

that I was thinking of places like downtown Bellevue or downtown Seattle, where there are quasi public spaces that are technically on private property, but they have provided access to the public, sometimes with limited hours, sometimes open 24 hours a day. And they're not necessarily maintained by the city, but they're a vital part of the urban environment for pedestrian access, and sometimes in downtown Seattle, where they're dramatic inclines, they actually provide escalator or elevator services, which are, I think, potentially really useful for many hours of the day. So I'm just wondering, when you have things like walk sheds, and I'm looking at the downtown Bellevue portion of the map. It seems like some really, really key pieces are not there, but they're also like, I think they fall under the quasi public infrastructure, but would be really, really useful to kind of be included. So I wasn't sure if there was, like, already some kind of thinking behind what would be included in this database and what would be excluded.

Anat Caspi 49:44

So the official scope of this project is around roadside infrastructure that is publicly maintained, at least that was kind of the understanding and agreement, and somewhere we named that as part of Project. That is not to say that that really demotes the need and usability of the including the digital account of the infrastructure that you're talking about, I mean, access maps. You know, glory in Seattle is about including those indoor elevators that provide so much useful you know, extension of walk sheds in the downtown area. It's just not part of this collection. But the tooling will provide will enable people to extend the data to include those, and actually in Bellevue, I believe are because so much of it has been included in OpenStreetMap, I believe we can incorporate that as well. So does that answer the question? Yes, absolutely.

51:02

Thank you. So

Anat Caspi 51:03

essentially, what that means is, once you the data is is published for you to edit, or everybody to edit, really, and you want to become like a steward of a specific data release you can and the transportation data equity infrastructure allows you to basically grab a data set and say, you know, I have a next release, because I'm going to throw in all of these trails and additional data and incorporate it into the next release of the data that I'd like to publish. And here it is. And so the tooling will allow you to add those for sure,

51:48

got it maybe just a extra question on top of that, I think I was zooming in, particularly to some transit areas. So I noticed, for example, the Bellevue transit center. It has some elements of urban design that are not, perhaps conforming, particularly with respect to what we would consider a crosswalk, but they're clearly kind of things with traffic signals and safe passage for pedestrians, they're more like painted differently with textures. And I'm kind of curious, like, are there specific definitions of infrastructure, like crosswalk, where that it's like pretty strict, or is that also just completely up for editing, particularly in the inventory that you're keeping, but also, obviously the extension, people can kind of like change that as needed. But I was really, really surprised to kind of see that, you know, one of the most pedestrian friendly places around Bellevue, there are just some key elements that kind of like contrast with what's on the ground.

Anat Caspi 53:00

Um, that is a really good question. So you're talking about, like, active transportation features that have been added in but are not, maybe non standard, so they wouldn't be potentially recognized by the vision algorithms, right? Because they're not standard, yeah, so there's no, like,

53:20

zebra or parallel lines that denote a crosswalk, but it's the entire thing. It's painted red and has, you know, signals that the vision wouldn't catch, so it wouldn't be standardized. And I'm not even sure if it falls under some green book category of crosswalk or not, but it's very clearly used as pedestrian facilities. You know, there, there are very few of these, I think. But I wouldn't surprise me if we kind of, like get into territories of some jurisdictions where experimentations are being done, like that. Yeah,

ద 53:58

no, for sure. So I mean, this kind of falls into the realm of the local deep dive because because of its regional flavoring, but I think it's important to talk about that for sure. I think again so as stakeholders take on stewardship for a particular data set, and want to include that. I think it would be great to do that. It's just outside of our sort of you know, what we're trying to do is the most consistent, uniform, high quality data we can provide. And so unfortunately, that means we can't do one offs all the time. But we we do want to know where our algorithms fail for sure if it's something that we can actually correct without including, without bringing into any false positives, because, like any red marking, obviously, is not something that we want to include. If that makes sense, I hands

A   Anat Caspi 55:03

anyway, if you can send us pictures, that would be fantastic, like of the area that you're talking about. All right, am I missing hands because I can't see much cool. All right, so moving on our complete QC analysis pipeline. So what has happened so far? I talked a little bit about this, but we should be really clear. So we took multiple sources, satellite imagery, the OSM road network and crowdsource data. We created computer vision pipeline to predict where these features are that we talked about. So sidewalks, the links from the sidewalk to the crossing and the crossings and we provide the metadata documentation to ensure that all the data sources are documented with the source the data of the collection and different confidence metrics that we have. And in addition, we do an insult called validation to just perform the basic check for data completeness and format consistency, and, you know, plausibility, essentially using some scripting. The next step is then to provide remote verification through mappers who want look at the satellite imagery from which the data was predicted and validate the critical nodes, the paths and for all the predictions. But what we're hoping to work on with you is this kind of expert review assemble information about where the data does not seem to have high confidence or high quality, especially in areas that might have high discrepancies. Suburbs tend to be harder. Areas that have been traditionally industrial because of the and but are now transitioning into new residential zoning. A lot of times there's lots of discontinuities in sidewalks, because projects kind of require sidewalks as housing developments come into play, but they're not entire blocks at a time, so those are going to be the areas to kind of watch out for and let us know if you see high discrepancy between what's on the ground and the data. And then we're also, as indicated, we are working with communities to get community feedback, to engage local communities and stakeholders to report any issues through the tooling platform. The follow up QCS really have to do with both automated quality checks, but also cross referencing with external data. So it's not going to be for all predictions. It's not as comprehensive as that, but for all of the jurisdictions, we've collected points of interest, like transit, like food and grocery stores, like schools, and we're basically doing scenario based evaluations for the walk sheds in those areas, using the walk sheds tools that you saw to understand whether the access and reach targets, these points of interest, have the kind of Data and connectivity that could be validated on the basis of other routers, but of course, other routers don't actually have the sidewalk infrastructure data, and so we're using essentially surrogates about routability based on some GPS traces that we might see from like Strava and other things like that, as well as other indicators that provide us with these scenario based evaluations or speculations about where the data might not be connected but should be. And then we're integrating with other authoritative data, like road networks, and also, as I mentioned, looking at zoning as predictors of where connectivity will be good or not so good, as well as changes in zoning, so where it used to be industrial, but changing to residential, we

anticipate still having pretty low connectivity. And so we're looking to see if the data sort of comports with our expectation on on connectivity. And then once we have data that we consider to be high confidence, it's fed back to the model so that we can improve over time. So that too speaks to the effort to make this a sustainable and an ongoing, updatable effort, as opposed to a one time deal that's just dumped out there and dead on arrival, so to speak, I saw something in chat was that, to me, are points of interest data available to agencies or it's not proprietary, and we are super happy to share what we have, and happy to have additional POIs indicated to us if they are high importance. In fact, like the more points of interest you share with us that might be, you know, areas of part of the h, i n, the or other areas of the network that might be needing additional evaluation or of concern, we'd love to hear about them. And again, a way to reach out to us would either be to just email, or once the commenting function is turned on in the viewer, you can just like comment on that and it would allow you to essentially send us an email about that specific location. So where are we going next? So we intend by the end of December, to complete 50% of the intended trajectory for the biennium and finish the three to five regions of Deep Dive. We continue to work on data review and best practices, and we hope to have an additional work group meeting before December, so that you can give us some feedback on the annual report, as well as some of the data review, if you get a chance to do that, and we'd be really excited to hear back, and that's when we'll start talking about the analytic workflows. So we've already talked to some of you, stakeholders@wash.at Sound Transit planners at King County, and basically the two main use cases that people floated as far as like analytic workflows of interest had to do with project planning. So being able to see to understand how access and walk sheds might be changed by a specific project on the ground. So basically, what if type of scenarios being able to look at specific sidewalk areas and say, Oh, if I put in these two curb ramps, how does that change access to in this area? And the other analytic workflow that people were requesting were essentially just routability, so from one location to another, but those were specific interviews that we worked on with accessibility folks, because that's one of our foci in this work, we're happy to hear about additional workflows that might be of interest. I know, for example, that Thomas Craig is very interested in integrating this into a multimodal network, and has been working with conveil on some of that workflow. So we're open to additional information about the kinds of workflows you might use or need this and utilize this data would be useful for, I think Justin indicated a little bit of that with analytics around Sound Transit facilities. Should we open it up for comment at this point? I think it's a good time to do that. So also coming up in that meeting, we'd love to get a little bit of review of our annual report, and we should have that out in the next few weeks. It might not have all the up to date information that will be the final data discussion at the end of December, but it should be sufficiently comprehensive to kind of indicate what it is we're doing and provide the state legislature with information about where data has been collected, how the data collection proceeded, who was involved in the procedure and in the process, et cetera. And so towards the June deadline, we intend to complete the full QC pipeline, publish the data and analytic pipelines as well as the data, and continue to work on additional QC and metrics, because that's really important to do, and we hope to engage you in our continuing work. So before we close, well, for one, I do want to question whether there are specific analytic workflow interests that are out there and that you'd want to work with us on. I lost

1:05:27
some folks,

Anat Caspi 1:05:32

all right, and so preparing for next meeting, we will publish the Publish to you the annual report. You stakeholders who have involved, been involved with us, we're still interested to hear you folks out there in Spokane, we'd love to hear from you on as priority regions of interest for additional deep dives, the Bellevue concerns on like specific regional features was also interesting in this context, and then again, clarifying the analytic paths, meaning any specific workflows that you can envision using this data for and how we can better integrate this work into the workflows that are already used in your institution, already used@wash.we have a specific project that we'll be launching in then the next couple of weeks, really urgently actually, to start working with stakeholders At wash dot, specifically to understand that kind of integration. So we hope to kind of prime ourselves up for for that, as well as additional workflows that might be of interest to other institutions as well. Any comments on that? Did I lose everyone.

1:07:07

I think you got a comment,

1:07:10

yeah. So this is Yes, Kevin city, so can so you know, this is our first time joining this meeting. I guess I'll just share and ask if you're hearing a similar thing from other agencies, or if you had a similar question. You know you've noted that there's regional data that we have for sidewalks, and so it's decent data's maybe not perfect, and maybe could use some cleaning up, but we do have something our currently, our and lots of discussion going on internally on this, our bigger area of data need, and really a whole is, is sidewalk condition data, because we were, we're a fairly old city, you know, we were roughly 100,000 people just after the turn of the century in 1905 so you can imagine, 40% of our city, roughly, that exists today was, is 120 something years old, and we got a lot of neighborhoods with big, big trees uprooted sidewalks and some very severe sidewalk challenges. So that's kind of our bigger challenge right now. Yes, there's areas of missing sidewalk and missing ADA ramps, and that's important, and that's definitely something we're working on, but that sidewalk condition data is something we're talking about and struggling with, how to how to deal with capturing that. So if you had a similar question come up from other agencies, or any dialog on

Anat Caspi 1:08:45

that, yeah, completely. How to make this succinct? Because we don't have all night. So data quality, or rather a sidewalk quality is difficult because there are so many variables we've been really focused in this effort. Obviously, like the sheer scale prohibits us from being on every sidewalk everywhere, so it makes it really difficult to do the on location quality assessment. And you know, satellite imagery doesn't quite give you that the state is collecting LIDAR data for state controlled infrastructure. And so we have, in the past, you know, done some prototyping on how to glean that information in a very consistent way from from the LIDAR imagery. It's not part of this project. However, what we're really aiming for is the ability for different stakeholders at whatever level to be able to collect this data at, you know, whatever

tooling is available to you. So if you have LIDAR information, and you have point location information about the surface disruption, the quarter inch displacement, etc, you can join that data with this data, so that you'd still have the baseline open sidewalks data, but have those additional quality information added to that data set. In addition to that, for example, the King County Metro local dive that I talked about, they're interested in quality assessments, but they do so on a per sidewalk basis, and so we with the app that we publish for their use for that extension, the additional attributes each and every sidewalk or sidewalk segment is assessed for, like how many surface disruptions did you see in this segment of the sidewalk? And so you can just count up the number. It's not like, you know, high resolution quarter inch displacement data, but it gives them enough information to assess. Like, should we offer, you know, in our next project, you know, should we prioritize this piece of infrastructure over another on the basis of, like, how many disruptions per 100 feet of infrastructure, for example? So in the app, you're able to just go and collect that. And so the Go info game app, which I mentioned earlier, is part of this local dive, but it will continue to exist. And then each institution can kind of identify, you know, what is the form? What is, what are the set of questions that you wish to answer, to have answered on every piece of infrastructure. And it tracks, you know, whether the data is complete, whether the tags were added for all of the infrastructure pieces. We kind of gamified so that we're also able to to publish it to crowds. So something that started yesterday is this team challenge for just people on the ground who want to do data collection, so there are different teams who are able to use the app and make the assessments. And of course, we're not immediately incorporating random crowd sources data into this data, but it's a way to get people both understanding that this project exists, that this data is being collected, that you know, increasing public awareness for like access and reachability concerns and things like that. But the point is that we're making this really accessible so that, but you can still ask a very detailed, high resolution question with those quests like requiring a meter stick, let's say to ask, what is the remaining sidewalk beyond this disruption in the surface that allows people passage here, which is another way that people have opted to collect data now, when it comes to access map and open sidewalks, our preference is to look at data that we consider to be neutral and objective, and of course, those are both kind of frayed terms, but the idea Being that the collector is not biasing the data. That's an outcome here. So what we ask people to assess when we do our own like open sidewalks extension, it's like, if there's a disruption in the surface, is it a vertical gap and is it horizontal? So that's something that people would be able to assess. And then what is the usable remaining widths of the path so that we are able to then simulate, you know, what would happen with a wheelchair of standard size, what would happen with a car, what would happen With a bike, etc. So again, the tooling is

1:14:23

agnostic as to what it is that you're collecting. It's kind of about the handshake between whether your collectors have the capacity to do the collection of what you're asking them to but the tool would enable you to create your own form over whatever collection, however you decide. Now, the back story to all of this is that in 2018 Seattle had a lawsuit with a with a settlement that allowed them to provided the funding for 19 U DUB interns to do a collection on data quality, on surface quality in sidewalks in Seattle. And those interns walked every inch of Seattle sidewalks, but the way in which the data was collected ended up really enriched for whoever was collecting the data, either it was a flaw in training or a potential flaw in having too many overlapping definitions of like quality concerns. So somebody might call this a vertical disruption. Other might just call it a gap, but both of those were available as annotations, and so we really have not been able to use that data effectively to report to our Access Map

stakeholders. What is the true quality of this data? Like it was not, or true quality of this infrastructure, we've not been able to provide good enough information for somebody to be able to interpret back from the app, like, oh, would this be good for me or not good for me? Can I will it be bumpy here or not? So? So we've opted to not use that data, and that's too, too sad, but it really motivated us to provide the best tooling possible for collecting things that would be interpretable both to programmers on the other end, but also to folks on the ground.

A Anat Caspi 1:16:27

That was a very long conversation. This answer, I hope it answered some of what you were asking.

1:16:35

Yes. Thank you.

1:16:38

What I can give you. So in our most recent Open sidewalks advisory group meeting, we did have, like a group of 30 folks who essentially voted in what they thought would be the most effective way to collect this information, modulo quarter inch displacement data. So if you're if you don't have access to LiDAR, you know, what's the best way to collect this information that would be interpretable? And basically the outcome was, like, this hierarchical approach. One is to say, like, you know, is this a vertical or horizontal gap? The other was to just measure the impacted area, so, like, what's the width and the height of this area and what's the effective pass remaining?

A Anat Caspi 1:17:24

And then some additional features about greats and gaps and things like that. And I'm happy to share that with anyone who cares for it.

T thomas craig 1:17:40

I'm wondering this conversation is bringing up for me, because I don't know, you know Kevin, whether I guess it sounds like it would be fruitful to hear from Spokane, kind of like what you'd want to do with quality data, and like what granularity you imagine that data being, you know, feasibly used by, you know, for your business process on because I'm also wondering, you know, there's, there's ideal scenarios of the data where we're really getting into, you know, modeling gaps. There's also, you know, kind of practical use cases around, I mean, like, maybe it's just, if we're, if we are talking about Asset Modeling or ask, you know, figuring out planning for spending on various physical assets. Maybe it's just a matter of, you know, this segment has a must rebuild by x date tag on it, right? You know, I don't know what that looks like from the

perspective of, you know, regional or city planning, but maybe it's a much simpler data model. So I'm curious, you know, kind of what, what the business use case is, Kevin, and if you you know, maybe that's for another conversation. But I put the question out there,

1:19:01

no, we love to hear that, actually, yeah, again, just share the issue and concern and desire is that is to is to start to address our our port, our significant amount of sidewalk And poor condition, again, mostly age sidewalk up, you know, tree root uplifts, those sorts of things. And we don't have a anecdotally, we kind of understand where our problem or challenging neighborhoods are, in general, but we don't have that comprehensive data to really know, and not granular from my end, as more of an engineer, maybe being too specific. But, you know, it's not as granular enough. You know, if someone asked me, What is our need citywide? What would that cost to go replace every stick of, you know, a non Ada, accessible, uplifted, poor condition sidewalk? I have no way to answer that question, right? And it gets pretty granular in terms when you're going into these built environments of, well, do I need to replace one panel a sidewalk? Is it two panels a sidewalk? You know, those details, we have some anecdotal information. But without those assessments, you know, you're I'm fearful that the data is not, you know, going to be accurate enough, even if someone took a wagon. So are you? We're just grappling with that. We don't have a great funding source, frankly. So even if we come up with all that data, they're still grappling with the funding side. And just the challenge that everyone every agency, has is that technically, sidewalks are the responsibility of adjacent property owners, you got affordability issues of, you know, telling someone they have to for code compliance, they have to go replace their poor condition sidewalk when they can't afford that, you know. So the whole funding part of it is a whole other animal that's almost a moot point until someone figures out how to fund that need, because our just our routine maintenance sources are would just be a drop in the bucket. Wouldn't be substantial addressing the problem,

thomas craig 1:21:12

I'll say, maybe briefly on that topic. I think one of the potential values that most valuable things that comes out of this project is being able to quantify that unbudgeted cost, and you know, present what order of magnitude it's on, with some detailed backing, scientific backing, because it is significant and will ultimately need to be accounted for,

Anat Caspi 1:21:38

I can share so we had one student group do that kind of assessment for Seattle on the basis of that, even, you know, the data set that was concerning, just to do that simulation, you know, on the basis of, we kind of took the data set, we gross define what which issues might be addressed, and then Seattle, D O T provided us with kind of a price menu of what it would cost to improve those different types of concerns or surface disruptions. And so we came up with a cost model that was able to make an assessment for walk shed specifically. So, like, if you would click on the tool and say, you know, what would it take to fix the 10 minute walk shed for this bus stop, then it would, you know, provide that kind of assessment on the basis of aggregate. So it's a very gross tool. I don't think it would be, but it would still provide a way to

prioritize certain areas versus others, potentially, if you believe the underlying surface disruption data, thank you. So this pretty much was everything we wanted to report with the call to action to get more of your feedback. We're really interested in what you have to say. I hope it's clear that you know, this project wouldn't have been possible without the feedback that we've gotten throughout, specifically with Thomas's involvement, but also a lot of info from EDA from Grace, from Mary Elizabeth was helpful, and the King County team from health through housing. So if we're available for conversation, if you have time to talk to us, and I invite you to reach out. All right, any let Yes, go ahead. Justin,

1:24:14

I was gonna say thank you. I have to drop off at 430 but this was always helpful, and I have some things I want to bring back to Sound Transit to see if some additional feedback for you.

A Anat Caspi 1:24:25

Awesome. Thank you so much. And thank you everybody for joining us again. Thank you for action.

1:24:34

Thank you. Have a good one.

1:24:39

Thank you. Applause.