OpenIR [Open Infrared]

Enhancing Environmental Monitoring Through Accessible Remote Sensing, In Indonesia and Beyond

Arlene Ducao

M.F.A, School of Visual Arts, 2004
B.S. University of Maryland, 2002
B.M. University of Maryland, 2002

Submitted to the
Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology

June 2013

Author_______________________________
Arlene Ducao
Program in Media Arts and Sciences
May 10, 2013

Certified by_______________________________
Henry Holtzman
Research Scientist
MIT Media Lab

Accepted by_______________________________
Prof. Patricia Maes
Associate Academic Head
Program in Media Arts and Sciences

OpenIR [Open Infrared]

Enhancing Environmental Monitoring Through Accessible Remote Sensing, In Indonesia and Beyond

Arlene Ducao

Submitted to the
Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology

Abstract

As the human landscape changes ever more rapidly, environmental change accelerates. Much environmental information is publicly available as infrared satellite data. However, for the general user, this information is difficult to obtain, and even more difficult to interpret. With this in mind, my team and I launched OpenIR (Open Infrared), an ICT (Information Communication Technology) that provides geo-located IR (infrared) satellite data as on-demand map layers, automates environmental feature classification, experiments with flood risk mapping, and interfaces IR data with crowd- and citizen-maps. OpenIR's initial use case is emergency management and environmental monitoring in the economically developing and ecologically vulnerable archipelago of Indonesia, where we conduced initial usability tests in January 2013.

Thesis Supervisor: Henry Holtzman
Title: Principal Research Scientist

OpenIR [Open Infrared]

Enhancing Environmental Monitoring Through Accessible Remote Sensing, In Indonesia and Beyond

Arlene Ducao

The following people served as readers for this thesis:

Ethan Zuckerman, Principal Research Scientist, MIT Media Lab [Reader]

Steve Chan, Ph.D., Director, IBM Network Science Research Center [Reader]

Contents

Chapter 1: Acknowledgements

OpenIR could not exist without the dedication and work of my core collaborators: Ilias Koen (The DuKode Studio), Juhee Bae (MIT), and Barry Beagen (MIT). Important MIT contributors also include Aziz Alghunaim, Stephanie, New, Wendy Cheang, Nori Hirashima, Srinidhi Viswanathan. Thanks also to high school intern Faisal AlHumaidan.

When environmental journalist Harry Surjadi became an OpenIR community partner, his immediate understanding of the long-term applications of the software helped to legitimize the project in its initial study area of Indonesia. Harry and Barry Beagen, both Indonesian natives, helped OpenIR's January 2013 fieldwork to be truly relevant and useful to all parties.

Deepest thanks to my thesis readers, who have also been great advisers: Henry Holtzman (Principal Investigator of Information Ecology, my research group at the Media Lab), Ethan Zuckerman of the Center for Civic Media, and Dr. Steve Chan of Harvard/IBM/MIT. OpenIR advisers Lela Prashad and Heidi Meisenkothen have also provided wonderful guidance. Thanks also to new OpenIR affiliate, Miho Mazereeuw of the MIT Urban Risk Lab.

In Indonesia, I very much thank those who took the time to work with my collaborators and me: Kate Chapman and Humanitarian OpenStreetMap Team, Katrina Engelsted, Fadhilah Mathar and the Indonesian Ministry of Communication's ICT Training Center, Annisa Fajriyah, Elvina Anita, Tanti Liesman and UN Global Pulse in Jakarta, Stefanus Masiun and the Ruai TV staff, Trias Aditya and Universitas Gadjah Mada (UGM) Department of Geodetic Engineering, Heri Sutanta (UGM), Rindi and Dinih (students at UGM), Emmanuel Migo and the Sekber REDD+ staff, Alfanius Rinting and the AMAN Kalteng staff, Ade Tanesia and the Combine Research Institute staff, John Taylor, Elisa and Edwin Sutanudjaja, Marco Kusumawijaya, Wahyu (Komang) Dhyatmika and Tempo.co, Kristy Van Putten and AIFDR/BNBP, Kaka Prakasa, and June Reyes. If I've left anyone off this list, please forgive me.

Thanks to our technical adviser Dave Thau at Google Earth Engine, as well as Evan Shulman at Google.com.

In its first months at MIT, OpenIR project started to accelerate due to the advice and support of these advisers and colleagues: Miguel Luengo-Aroz, Anoush Tatevossian, Robert Kirkpatrick, Heather Leson, Willow Brugh and Geeks without Bounds, Liz Barry, Leif Percifield, Jeff Warren, Juliana Rotich, Becky Hurwitz and the Civic Maps team at MIT, Ted Okada and FEMA, Desiree Matel-Anderson, Jesse Rozelle, Scott Shoup, John Crowley, Kristyn Cook-Turner, and Saurabh Channan.

The MIT Public Service Center (PSC) was a foundational element for the OpenIR project, and is the source of much of OpenIR's funding for Indonesia. Kate Mytty, coordinator of the MIT IDEAS Global Challenge, and Alison Hynd, coordinator of PSC Fellowships, have talked my team and me through much of our development process. Laura Sampath of the MIT International Development Initiative has also been a great supporter.

OpenIR was also supported by the South by Southwest Interactive Festival and by O'Reilly publishing. Special thanks to Hugh Forrest and Sara Winge from these organizations. Thanks to John Bracken and Chris Sopher of the John S. and James L. Knight Foundation for featuring OpenIR on the News Challenge site.

I don't know if anyone thinks the MIT Media Lab is an easy place. It's definitely one of the most exciting, challenging places I've ever worked, full of exciting and challenging people. In addition to the aforementioned Media Labbers, I'd like to thank David Hill, Ricarose Roque, and Jie Qi, who helped me prepare for Crit Day and other days. Thanks to Topper Carew for being cool. Thanks also to Info Eco, my cohort ("New Grads 2013"), Pattie Maes, Chris Schmandt, and Linda Peterson for feedback on my thesis proposal. Cesar Hidalgo's "Data-Centric Projects" and Ethan Zuckerman's "News in the Age of Participatory Media" classes allowed for major progress with this project. Joost Bonsen, the Media Lab's Ventures Guru, provided great suggestions through the Media Lab and the PSC. Thanks to Joi Ito for connecting OpenIR to parts of his network, and to Stacie Slotnick for highlighting the project in many ways, on many days.

Thanks to all of the interviewees for OpenIR's evaluations; their feedback is discussed throughout this thesis. Thanks also to all OpenIR survey respondents.

In recent months, a number of fascinating individuals approached the OpenIR team for future collaboration. They are too numerous to mention here, but I am very grateful for every message, conversation, and bit of contact with all of these folks.

Thank you to my family, the Ducaos in Maryland: Amon, Ayn, Amon II, and especially my mom Agnes, who proofread this thesis. Thanks also to my Baltimore best friend, Alex "Xander Galexy" Wright, and to my New York best friend, Crista Grauer. Even though we run the gamut of skin tones, you are my family, too. Bogota reunion!

Thanks to the whole NYC and Bmore contingent, including the Soho Seniors, the AMNH disillusioned, the SVA disbanded, the NYU crew, GSS in G, Marcie Prime, Spanish speakers like Maria B and Mike B, C-town refugees, and the Vega-Stace clan.

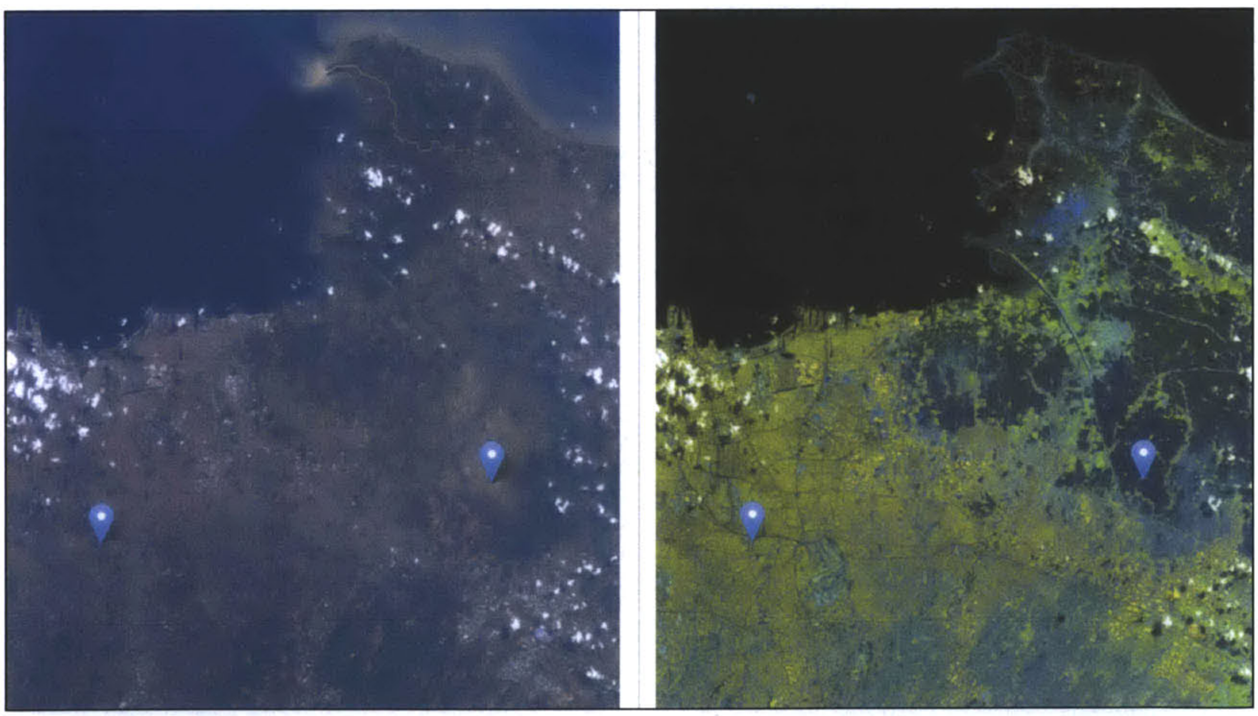
To my partner Chris Willard, for bearing with me through this whole Cambridge ("C-town") situation. Love for you and Brooklyn keep me coming back to the hood, coming back home.

Chapter 2: Background & Overview

Problem

When ecological disasters affect economically developing countries, affected citizens often lack access to the infrastructural systems that help save lives in more developed regions. For instance, with the massive 2004 Indian Ocean tsunami, economically developing, ecologically vulnerable areas were disproportionately affected, and almost 70% of the tsunami's deaths were in Indonesia.[1]

A large body of environmental information is particularly relevant to these situations: satellite data from the non-visible spectrum—specifically IR (infrared)—has been collected for decades by agencies like NASA, ESA, and JASA .[2] IR data allows for geospatial classification of environmental attributes like soil composition, building density, ground elevation, and water depth. This makes IR data useful for monitoring and predicting ecological phenomena like flooding and deforestation.



Comparison of true-color (left) and IR (right) satellite images of Jakarta. The IR image more clearly shows the difference between urban and irrigated areas (marked with the blue pointers).

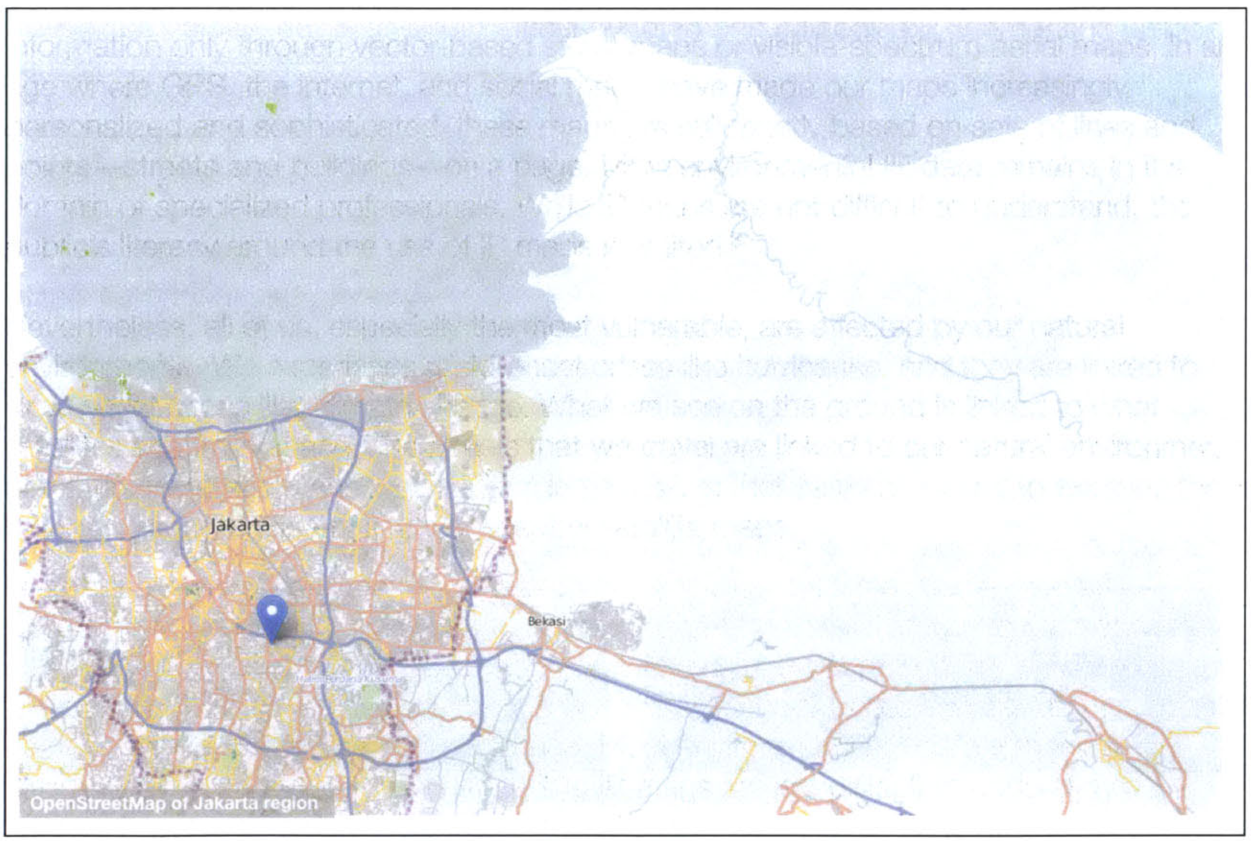
---

[1] "Tsunami death toll tops 118,000." http://www.cnn.com/2004/WORLD/asiapcf/12/30/asia.quake/

[2] Laken, Benjamin, Enric Pallé, Hiroko Miyahara, 2012: A Decade of the Moderate Resolution Imaging Spectroradiometer: Is a Solar–Cloud Link Detectable?. J. Climate, 25, 4430–4440.

While there is a large body of free IR satellite data online, it is not easy to use, particularly as free tools in web browsers. Our standard online map systems usually convey information only through vector-based street maps or visible-spectrum aerial maps. In an age where GPS, the internet, and social media have made our maps increasingly personalized and sophisticated, these maps are still mostly based on sets of lines and points—streets and buildings—on a page. Most environmental IR data remains in the domain of specialized professionals. While IR maps are not difficult to understand, the public's literacy around the use of IR maps is limited.[3]

Nevertheless, all of us, especially the most vulnerable, are affected by our natural environments. We experience rapid-onset crises like hurricanes, and they are linked to slow-onset crises like climate change. What we see on the ground is linked to what satellites see from space. The streets that we travel are linked to our natural environments of terrain, vegetation, and water. There is no system that easily allows us to examine these linkages through different kinds of spectral satellite maps.



Of today's web maps, OpenStreetMaps offers the most extensively drawn map of Jakarta. This is likely due largely to Humanitarian OpenStreetMaps pilot project in Indonesia. Nevertheless, large areas outside of economically developing cites like Jakarta (see east of the city in this image) remain unmapped.

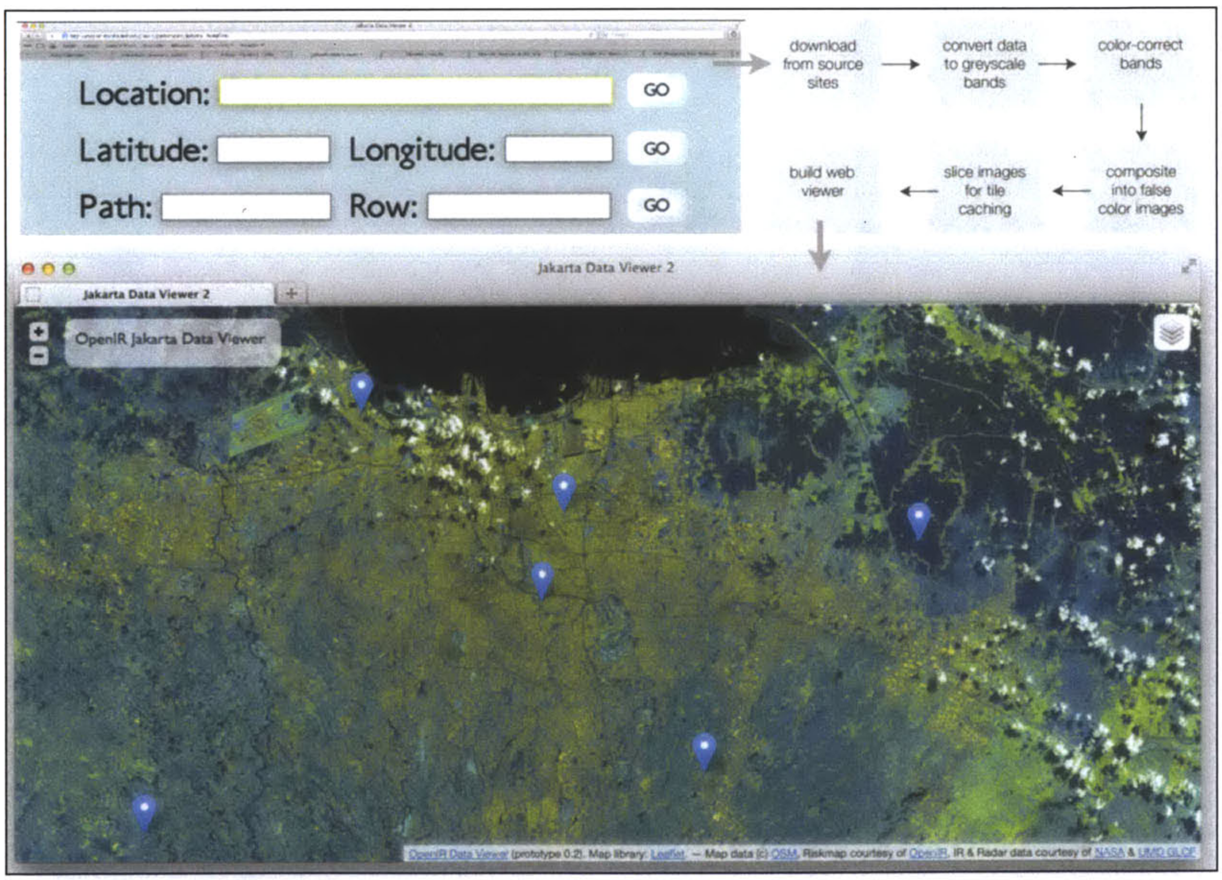
---

[3] Private conversation with Kate Chapman, June 2012.

Thesis Description

OpenIR (Open Infrared) is a free ICT (information communication technology) that offers geo-located infrared data as on-demand map layers, automates environmental feature classification, experiments with flood risk mapping, and interfaces all of these products with participatory maps. It supports efforts at ecological sustainability by highlighting environmental features not easily seen through street maps or true-color aerial maps. The United Nations Human Development Report 2011 defines sustainability as multi-generational equity over time: the level at which subsequent generations can access the (usually natural) resources of previous generations.[4]

OpenIR also explores the intersection of Pervasive and Participatory Sensing, a growing genre in Sustainable Human-Computer Interaction.[5] In this case, the remote (pervasive) sensing of earth-orbiting satellites is paired with the crowd (participatory) mapping of individuals on the ground.



Schematic of OpenIR's basic input/output system. A user will input a location, and the OpenIR system will output a data viewer of that location.

---

[4] United Nations: Human Development Report 2011, http://hdr.undp.org/en/reports/global/hdr2011/, 13.

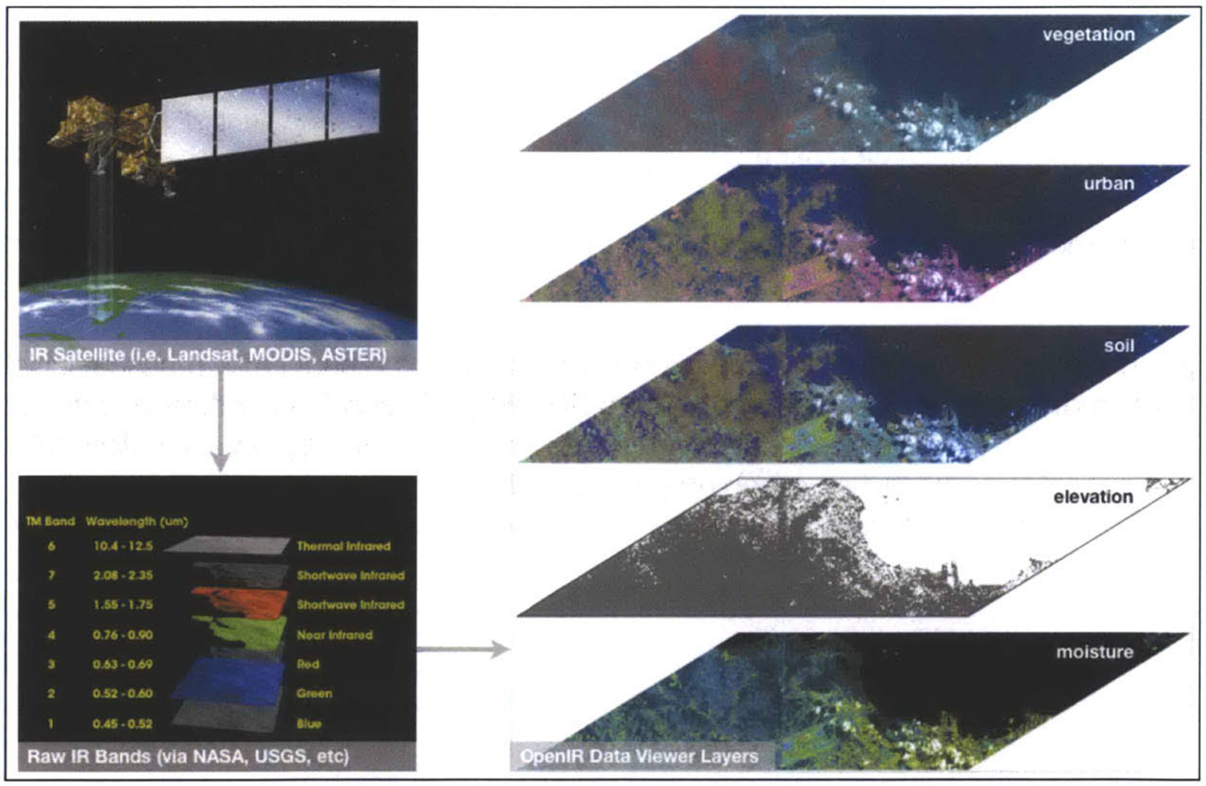
[5] Di Salvo, Carl et al. "Mapping the landscape of sustainable HCI." CHI '10: Proceedings of the 28th international conference on Human factors in computing systems, 1975-1984.

OpenIR's technical innovations include:

- Programmatic interpretation of IR data through eco-feature highlighting and automated risk map creation.
- Tools and plug-ins for general audiences to annotate and validate satellite imagery with on-ground observations.

Initial research with OpenIR examines how users regard trade-offs in spatial, temporal, and spectral resolution of IR data. No public IR dataset has high resolution in all three axes:

- Spatial: areas ranging from a single block to an entire continent.
- Spectral: ranges can be combined depending on desired environmental feature
- Temporal: Recent, historical, and (in the case of disaster) real-time imagery



Schematic of OpenIR's data pipeline and Data Viewer layers. Satellite & IR band imagery from Landsat. Technical implementation of this pipeline is explained in Chapter 4.

OpenIR research also includes experiments with different IR satellite datasets for environmental use with the general public: Landsat[6] and ASTER[7], which are commonly

[6] Landsat Missions. http://landsat.usgs.gov/

[7] ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer. http://asterweb.jpl.nasa.gov/

used for high-resolution environmental analysis,[8] as well as MODIS (high temporal resolution, low spatial resolution),[9] IKONOS (high spatial resolution),[10] and other IR satellite sources.

Concept image of IR data combined with participatory data (red circles). IR data can help highlight areas of refuge (pink outline) and vulnerability (blue outline). Combining maps of these areas with participatory data could help planners and responders better allocate resources for citizens during emergencies.

Through its user interface tools and its plugin for the crowd-mapping software Ushahidi, OpenIR will allow for research into patterns that may arise when participatory map data is combined with IR satellite map data. I hypothesize that combining these datasets and examining their patterns can greatly aid the planning, response, and analysis of ecological

---

[8] Joris Verbeken, Leen De Temmerman, Rudi Goossens, Philippe De Maeyer, J Lavreau. "Classification of the vegetation in the Virunga National Park (D.R. Congo) by integrating past mission reports into Landsat-TM and Terra Aster sensors." 01/2005; In proceeding of: Proceedings of the 24th EARSeL Symposium 'New strategies for European Remote Sensing', Dubrovnik (Croatia), 24-27 May 2004

[9] MODIS: Moderate Resolution Imaging Spectroradiometer. http://modis.gsfc.nasa.gov/

[10] IKONOS from GeoEye. http://www.satimagingcorp.com/satellite-sensors/ikonos.html

crises. However, while this thesis makes the combination of these datasets easier for the user, a robust test of this hypothesis is beyond the scope of this Master's thesis.

During the thesis period from January 2012 to March 2013, we were able to test whether users thought OpenIR data viewers conveyed IR information in an understandable, useable way. We were also were able to verify utility of this data in predicting flooding, though we were not able to robustly test the crowd-mapping or produce general interest maps for slow moving crises.

A final research question focuses on aggregated interpretive maps, i.e. risk index maps generated by the OpenIR software, which can be used for slow- and rapid-onset crises. What is the best way to auto-generate these maps for general audiences? Could these kinds of maps illustrate slow-onset crises like corporate deforestation and changing species diversity? Again, my aim with this thesis is to start building tools to help answer these questions.

Resources Required

OpenIR is set up on a Media Lab 300 GB virtual machine, http://openir.media.mit.edu. This machine is running the Ubuntu operating system, and has Apache, MySQL, PHP, and Python installed. It also is running PHPMyAdmin and ISPConfig.

OpenIR Data Viewers can be run in an updated web browser with Javascript activated. OpenIR's back-end source code, which is used to build the standalone Data Viewers, is on http://github.com/dukode. To run this code, these installations are required:
- Geospatial Data Abstraction Library (GDAL), a widely used raster map processing package [11]
- Libgeotiff, a library for processing GeoTIFF files. This library is bundled into some GDAL packages.
- ImageMagick, a command-line image processing package
- WGET, a GNU software package for retrieving large numbers of files via HTTP and FTP

OpenIR's Ushahidi plug-in requires the installation of Ushahidi; see http://ushahidi.org for system requirements.

The open-source package Quantum GIS[12] was used to conduct data validation for OpenIR's Risk Index Maps against observed data from the Jakarta 2007, Jakarta 2013, and Hurricane Sandy floods.

[11] GDAL - Geospatial Data Abstraction Library. http://www.gdal.org/

[12] Welcome to the QGIS Project. http://www.qgis.org/

Community, Validation, and Evaluation

OpenIR targets general audiences who have basic understandings of the internet and map reading. The application's early adopters are to be non-technical map users with an environmental focus, like urban analysts, governmental and non-governmental workers, and journalists. OpenIR's primary study area is Indonesia, where ecological and economic vulnerability is apparent from frequent seismic activity and limited supporting infrastructure. OpenIR also offers participatory map for New York City in the wake of Hurricane Sandy, and validations were conducted with observed data from both cities.

Early in the OpenIR project, local interest was expressed by key players in emergency management (FEMA, Harvard Humanitarian Initiative), science journalism (Reuters and BBC), insurance (State Farm), and scientists in the field of satellite and environmental imagery (UNOSAT, NiJel, NASA). I conducted Front-End interviews with many of these players during Implementation Phase 2 (Chapter 5).

OpenIR is at the stage where community needs must still be extensively evaluated. ICTD professionals point out that "financial sustainability must stem from community needs."[13] With support from an MIT IDEAS Global Challenge Award, OpenIR was evaluated, via interviews and "talk aloud" sessions,[14] throughout Indonesia in January of 2013. OpenIR's community partners for this evaluation were the United Nations Global Pulse Jakarta Lab, the Humanitarian OpenStreetMap Team's Indonesia project, the Republic of Indonesia's Ministry of Information and Communication Technology, Universitas Gadja Mada, Korupedia.org, and Knight International Journalism Fellow Harry Surjadi, who works to develop indigenous citizen journalism in Indonesia.

From June 2012 to March 2013, OpenIR underwent three rounds of human-centered evaluations. OpenIR's Risk Index for flooding underwent three rounds of data validation.

---

[13] Raftree, Linda (2012). "On Financial and Other Types of Sustainability in ICT4D Initiatives." http://www.ictworks.org/2012/03/30/financial-and-other-types-sustainability-ict4d-initiatives/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+Ictworks+%28ICTWorks%29

[14] Lewis, C. H. (1982). "Using the "Thinking Aloud" Method In Cognitive Interface Design (Technical report RC-9265)". IBM.

	When	Subject	Questions
Human-Centered: Front-End	June-November 2012	Interviews (n=10) Anonymous Surveys (n=30)	General questions, i.e. do you want to see infrared satellite data more accessible in web maps? Combined with participatory maps?
Data Validation I	November-December 2012	OpenIR maps vs. Jakarta 2007 flood extents	To what extent does OpenIR's generic flood risk map coincide with Jakarta 2007 data?
Human-Centered: Formative	November-December 2012	Interviews (n=10)	Discussion of OpenIR's usability, understandability, and usefulness
Human-Centered: Summative	January-March 2013 (launched in Indonesia)	Interviews (target n=20) Anonymous Surveys (target n=60) Software Analytics (event tracking)	In-depth discussion and survey of OpenIR's application to emergency management, forest monitoring, journalism, and civic mapping
Data Validation II	January 2013 (data collected in Indonesia)	OpenIR Risk Index Map vs Jakarta 2013 flooding levels (per neighborhood)	To what extent does OpenIR's generic flood risk map coincide with Jakarta 2013 flood levels?

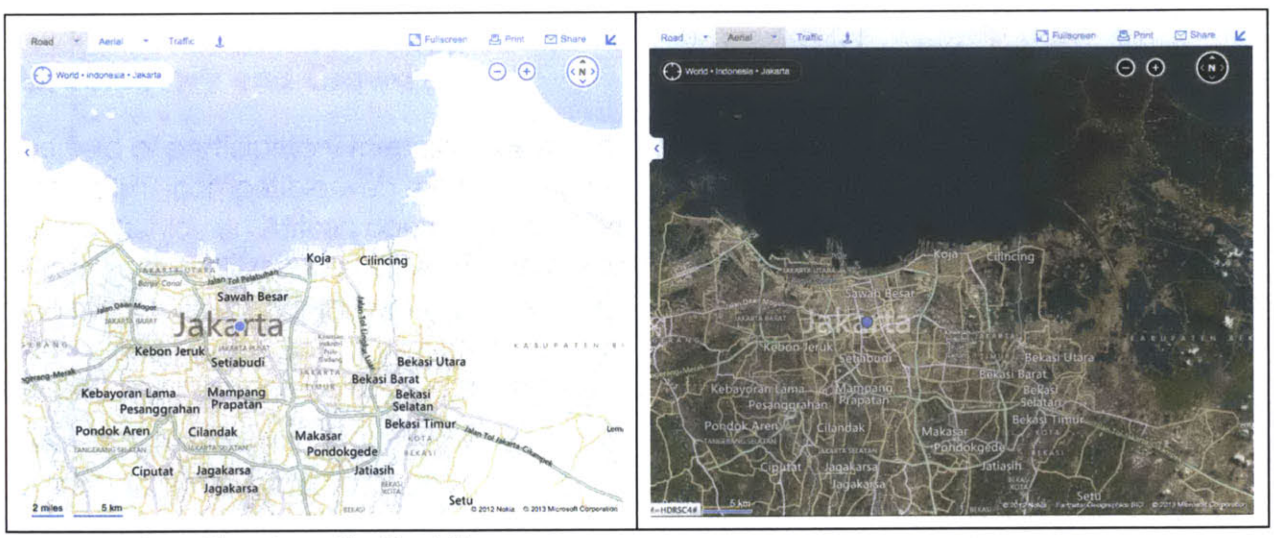
	When	Subject	Questions
Data Validation III	February-March 2013	OpenIR Risk Index Map NYC 2012 (Hurricane Sandy) flood extents	To what extent does OpenIR's generic flood risk map coincide with Hurricane Sandy flood extents?

Some of OpenIR's primary community partners in Indonesia. Work with Ruai TV and AMAN was organized through Harry Surjadi's previous fellowship work in Kalimantan (Borneo).

My long-term goal with OpenIR, via software development, community building, human-centered workshops, and data validation, is to make previously invisible information visible, thus empowering all of us, especially underrepresented and disadvantaged populations, with knowledge of the land on which we live.

Chapter 3: Related Work

From book-length street maps and atlases, to early web map ICTs (information communication technologies) leveraging the desktop publishing revolution of the late 1980s and early 1990s, to today's GPS-compatible (global positioning system) mobile devices, map ICTs have become increasingly more accessible, more up-to-date, more interactive, and easier to transport. Yet with all of the affordances of today's map ICTs, much of the information that they offer remains the same: vector-based points (buildings) and lines (streets), or raster-based, true-color aerial imagery. A comparison of the most widely used free web maps illustrates this point: Google Maps, Yahoo Maps, and Bing Maps all offer the same two basic layers, that of the street map and the true-color aerial map. While OpenStreetMaps, an open-source mapping application, offers a powerful participatory user interface, its information is limited to street maps.



Bing street ("road") and satellite ("aerial") maps.

Participatory and Crowd Maps

The field of participatory mapping has made great strides in the age of interactive, "real-time," GPS-compatible web maps. In addition to OpenStreetMap, "crowd maps" like those from Ushahidi, an African open-source software group, have been widely used for crisis response and citizen journalism during humanitarian and natural disasters[15]. However, as with most web maps, the default map layers in crowd maps are composed of vectors and true-color imagery. Spectral imagery taken outside of the true-color range, particularly infrared imagery, can be applied to crowd maps; but to date, this is an ad-hoc, poorly

[15] Ushahidi: http://ushahidi.com

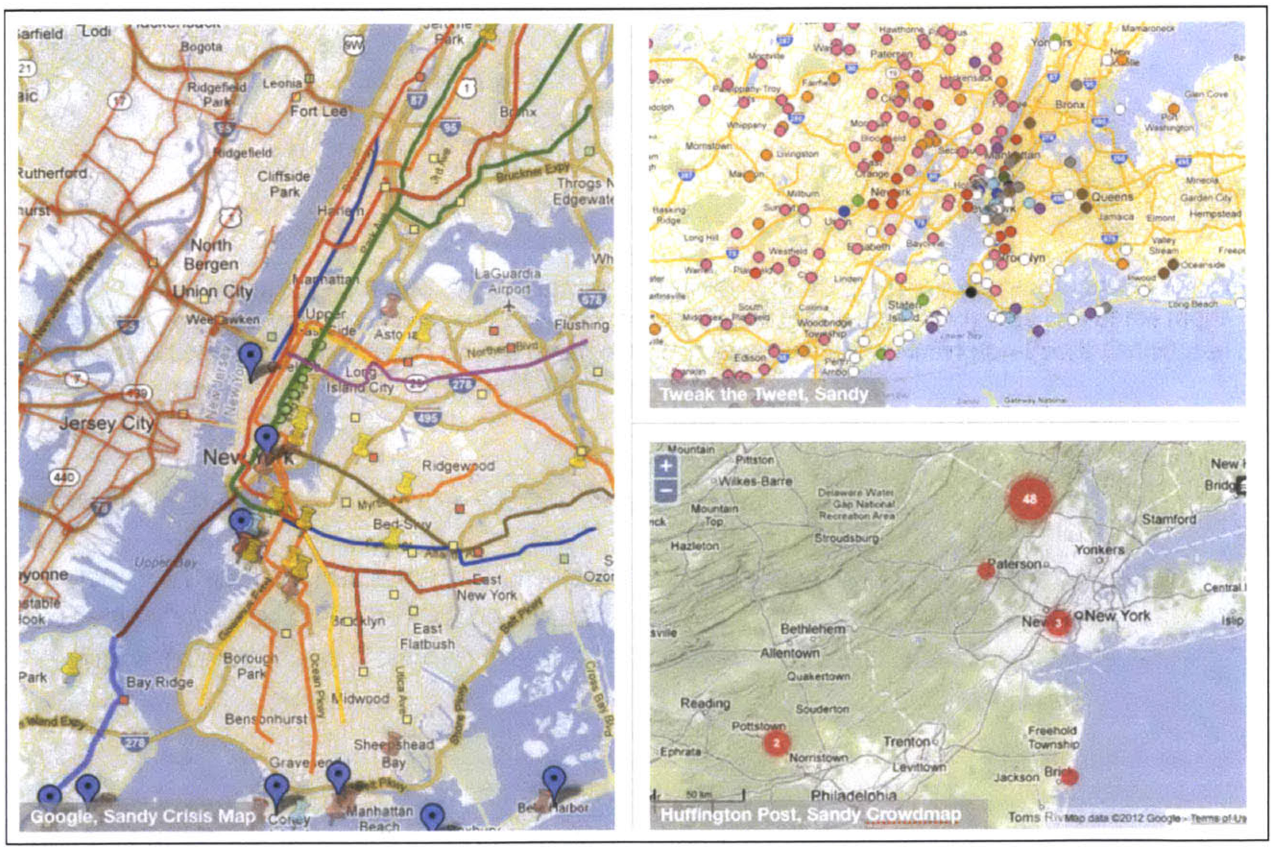
documented, and unsystematic process.

Ushahidi deployment originally initiated after the BP oil spill in 2010, by the Louisiana Bucket Brigade. The Public Laboratory of Open Science and Technology was also involved in this deployment.

Several crowd map systems, including Ushahidi, can input geographically located text from the platform Twitter. (Texts sent to Twitter are called "tweets.") One organized example is the "Tweak the Tweet" project from University of Washington, which uses a specialized, crisis-oriented taxonomy to better categorize tweets during a disaster.

The participatory maps used during recent crisis were perhaps a major influence in Google Crisis Response's decision to start building its own crisis maps, based more on municipal than participatory data.

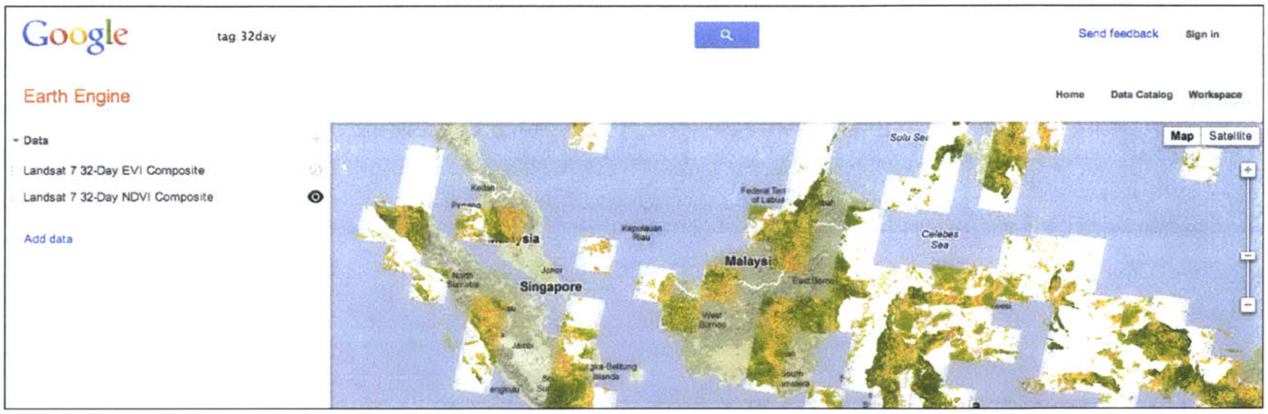
For all the rich information in these platforms, even in the case of environmental crisis, street maps and/or very simplified terrain maps were used as a base layer.



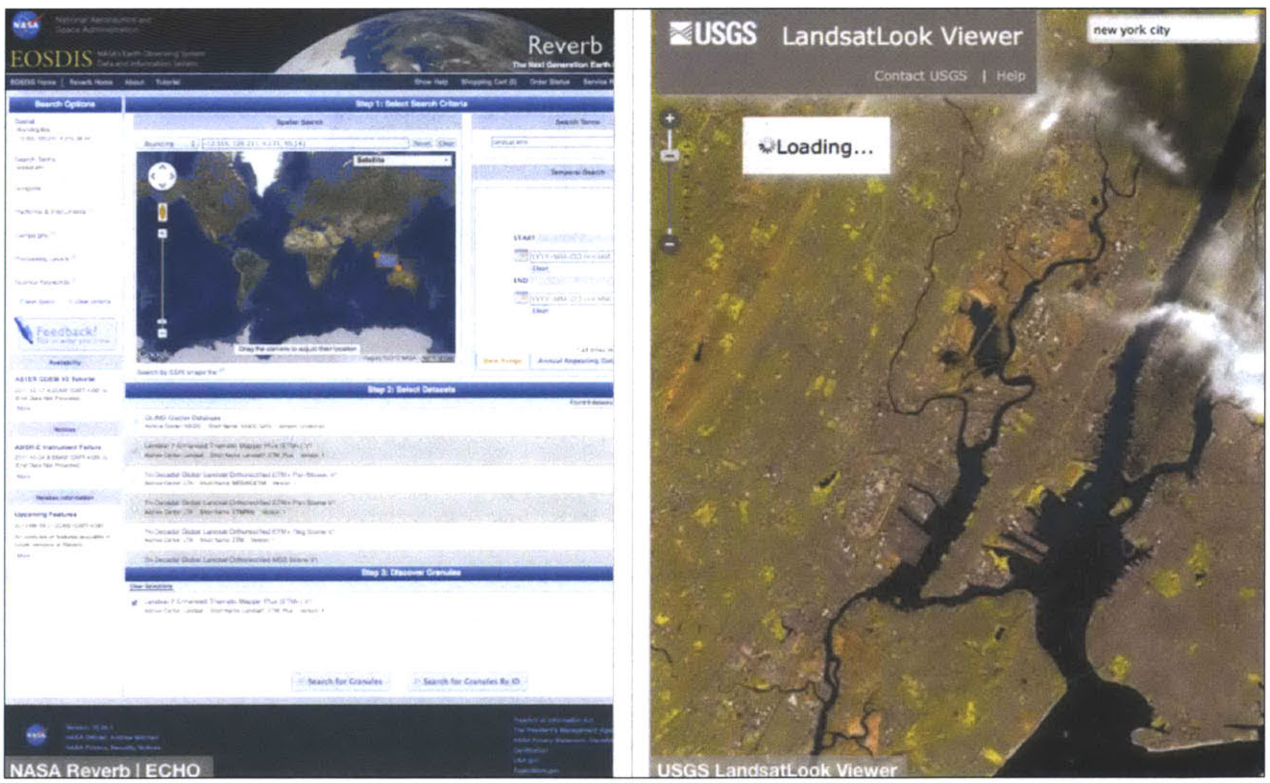
Popular maps used during Hurricane Sandy. Google's Sandy Crisis Map displayed municipal information like evacuation routes, subway disruption, and power outtages. Tweak the Tweet: Sandy was one of the most populated participatory maps, based on geo-located Twitter feeds, while the Huffington Post's Crowdmap deployment showed mainstream news media taking on participatory mapping.

Infrared and Spectral Maps on the Web

The use of IR satellite maps in web browsers is still uncommon, particularly because these datasets tend to be more multidimensional; they are collected in multiple spatial, temporal, and spectral resolutions. This can make the processing of big global IR datasets nontrivial, as can be seen in Google.org's Earth Engine: datasets are slow to load and often spatially incomplete. In addition, the tool targets environmental scientists and researchers, offering minimal guidance for general users.[16]



Google Earth Engine's "Workspace," showing Landsat NDVI (Normalized Difference Vegetation Index) composites for most land areas. Earth Engine is targeted to scientists and researchers; its API is not yet open to the public.



Two government portals for accessing IR satellite data. NASA Reverb | ECHO is the most extensive resource for downloading public satellite imagery, but it does not display the imagery in the web map. USGS LandsatLook Viewer displays IR imagery, but with little explanation or annotation.

[16] Google Earth Engine: Workspace. http://earthengine.google.org/#workspace

The same is true of institutional map portals such as NASA (National Aeronautics and Space Administration) Reverb ECHO,[17] USGS (United States Geological Survey) LandsatLook Viewer, and FEMA (Federal Emergency Management Administration) Map Service Center.[18] The user interfaces for many of these portals have either little interactivity and/or no imagery display.

Commercial or Free?

There are existing commercial tools, which may one day enter the pro bono domain. For instance, Google Earth, the 3D map visualization software, was originally a commercial tool developed by Keyhole Inc. before Google acquired the company in 2004, rebranded the software, and made a free version available.[19]

Google, Microsoft, and other commercial companies pay for satellite imagery that has high spatial resolution. However, imagery with high spectral resolution may be an untapped resource in environmental applications for the general public; movements toward open government and open data are demonstrating that there are insights to be gained from making existing public, free data more accessible. For instance, Data.gov, an initiative of President Barack Obama's administration, aims to "to improve access to Federal data and expand creative use of those data beyond the walls of government by encouraging innovative ideas (e.g., web applications)."[20]

Accessible Risk Maps

Environmental risk modeling is a complex process, involving many kinds of input, from land/air/sea features to socioeconomic data. Because risk modeling is very specific to a place, local risk model processes are rarely transposable to other locales.[21]

---

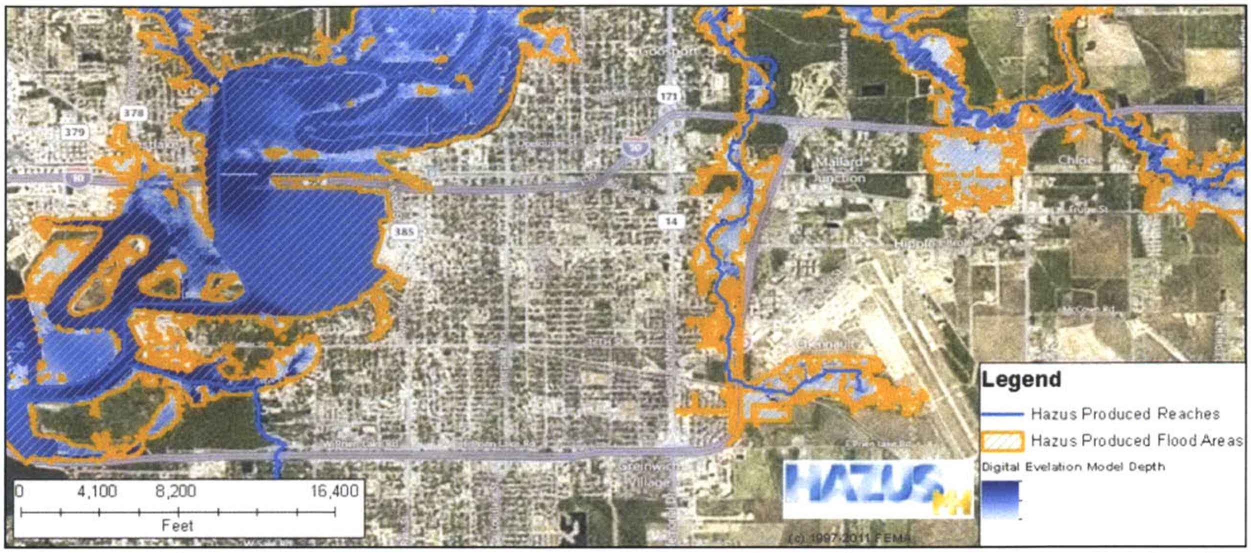
[17] NASA Reverb ECHO. http://reverb.echo.nasa.gov

[18] FEMA Map Service Center. https://msc.fema.gov/webapp/wcs/stores/servlet/FemaWelcomeView?storeId=10001&catalogId=10001&langId=-1

[19] "Geological Society of America 2010 President's Medal Presented to Keyhole Inc." http://www.geosociety.org/awards/10speeches/presMedal.htm

[20] Data.gov: About. http://www.data.gov/about

[21] Flood Risk Across Spatial Scales. http://www.understandrisk.org/group/drought-risk



FEMA's HAZUS software uses data collected and analyzed by US Army Corps of Engineers to produce flood insurance rate maps (FIRMS) for much of the United States. This software is free, but involves a large download, requires some training, and lacks a web-based interface.

Nevertheless, urban and crisis planners use more than complex risk models to organize community development. At the expert planning level, there are several free government tools to aid the process. FEMA's Hazus software[22] is one of the most comprehensive free risk planning tools available. It offers predicted flood extents for moderately damaging events (estimated for every "10-years"), as well as more damaging events (20-year, 50-year, 100-year, etc). These extents, available for some U.S. locales only, is based on data collected and analyzed by the U.S. Army Corps of Engineers. Since the original release of Hazus, a new software, Hazus-MH, shows risk related not only to floods, but to earthquakes and hurricanes as well. Neither Hazus nor Hazus-MH offers data that is directly viewable on the web, thus limiting accessibility to professional planners. While Hazus is free, they require Esri's ArcGIS, a high-cost software package[23], to run. Esri's Spatial Analyst[24] extension is additionally required for Hazus flood modeling functions to run.

AIFDR (Australia-Indonesia Facility for Disaster Reduction) has developed a simpler, lighter risk analysis software called InaSAFE[25] (Indonesia Scenario Assessment for Emergencies) as a plugin for the open-source software QGIS. For flood scenarios, InaSAFE uses observed data from the severe 2007 Jakarta flood, and then develops recommendations

---

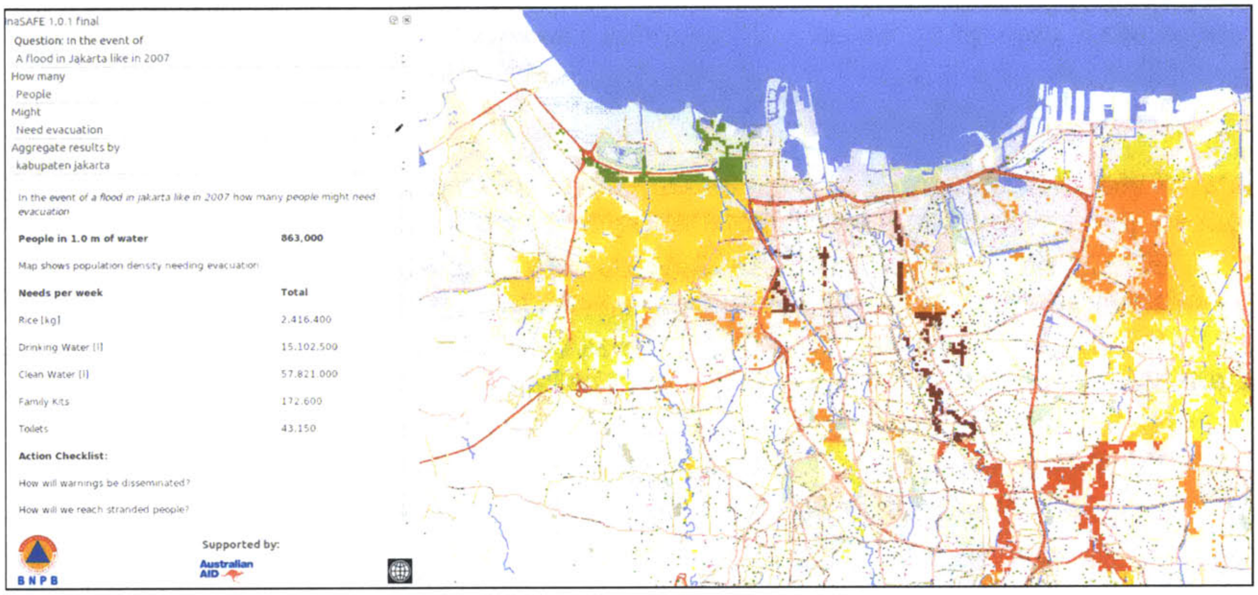
22 FEMA: Hazus. http://www.fema.gov/hazus

23 Esri ArcGIS. http://www.arcgis.com/about/

24 Esri Spatial Analyst for ArcGIS. http://www.esri.com/software/arcgis/extensions/spatialanalyst

25 InaSAFE Documentation. http://inasafe.org/

based on types of need: health, food, location, etc. Like Hazus, InaSAFE's data and derived products are not accessible via web interfaces.



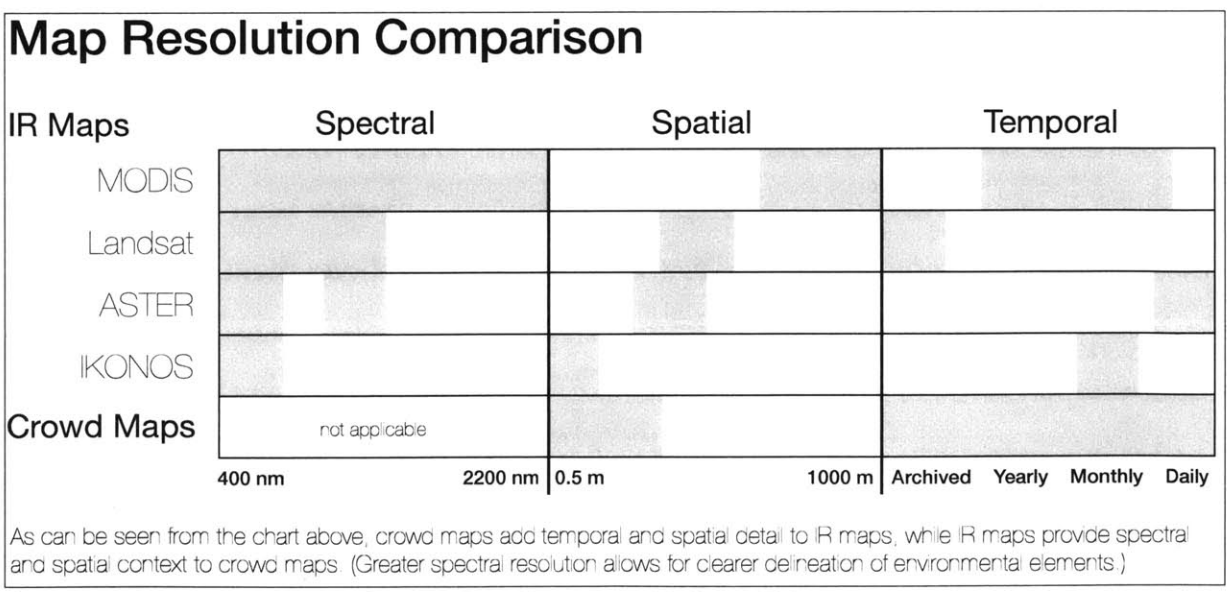
AIFDR's InaSAFE, a free plugin for the open-source software QGIS, is focused on Indonesia. It also requires some training and lacks a web-based interface.

Since a fundamental goal of OpenIR is greater accessibility to IR map data, initially for crisis applications, Hazus and InaSAFE play an important role. While these projects may be too complicated and/or early stage to be web-accessible at this time, they both provide comprehensive data and imagery that can be used for the OpenIR Risk Index validation process (see Chapter 10). There are also aspects from both Hazus and InaSafe risk assessment algorithms that can be applied to global IR datasets, and I incorporated these aspects into the OpenIR Risk Index algorithm (see Chapter 6). Finally, having talked to team members at both FEMA and AIFDR, I can see OpenIR playing a longer-term role in encouraging greater accessibility of both the Hazus and InaSAFE projects.

Chapter 4: Implementation Phase 1
January 2012 to May 2012

In OpenIR's first months, almost as much time was spent developing explanatory material (text, posters, movies, graphics) as implementing software or algorithms. Every winter, MIT's Public Service Center runs a competition, focused on funding international development and humanitarian work, called the MIT IDEAS Global Challenge. The process of entering this competition requires several rounds of concept refining, pitching, demonstrating, and community building. I entered OpenIR in this competition, not so much for the small cash prize, but more for the process of refining a concept for an international development project, particularly for a very broad project like OpenIR.

The initial process of implementation included
- determining IR layers for the OpenIR data viewer
- developing a simple web data viewer
- building an iOS prototype
- forming a team for the IDEAS competition
- iterating on concept and explanatory materials (i.e. posters, movie)
- transitioning to web-based prototype for IDEAS judging session
- manually building risk map
- building data viewers for other geographic study areas



Comparison of Public Infrared Satellite Data Sources, with Crowd Maps.

Preliminary Data Processing

Many visualizers of environmental maps use the University of Maryland's Global Landcover Facility (GLCF) as a source.[26] GLCF offers data from all three major IR satellite sources: MODIS, ASTER, and Landsat. Environmentalists working at the regional level and/or with many kinds of features tend to prefer Landsat, which offers good spatial and spectral resolution [see Ned Horning interview]. While GLCF's Landsat data is several years old, my collaborator Ilias Koen and I decided to use this data and data source for its ease (as opposed to the more difficult, comprehensive NASA Reverb ECHO) and the clarity of its imagery. We downloaded Landsat images for Jakarta, our primary study area, and New York, our home and secondary study area. Imagery from Landsat ETM+ (Enhanced Thematic Mapper) is divided into seven "bands" along the infrared spectrum:[27]

Band Number	Spectral Range (micrometers)	Type
7	2.08 - 2.35	Shortwave Infared (SWIR)
6	10.4 - 12.5	Thermal Infrared (TIR)
5	1.55 - 1.75	Shortwave Infared (SWIR)
4	0.75 - 0.90	Near Infrared (NIR)
3	0.63 - 0.69	Red
2	0.52 - 0.60	Green
1	0.45 - 0.52	Blue

These bands are commonly composited into "false color" imagery by assigning 3 bands in an RGB (red/green/blue) image, in which each band is assigned to a color channel. After consulting resources explaining commonly used band combinations, I chose the following band combinations to most clearly display the widest range of environmental features:[28]
- True Color: Bands 3/2/1 (all visible bands)
- Paved Surfaces: Bands 4/5/3
- Vegetation: Bands 4/3/2

---

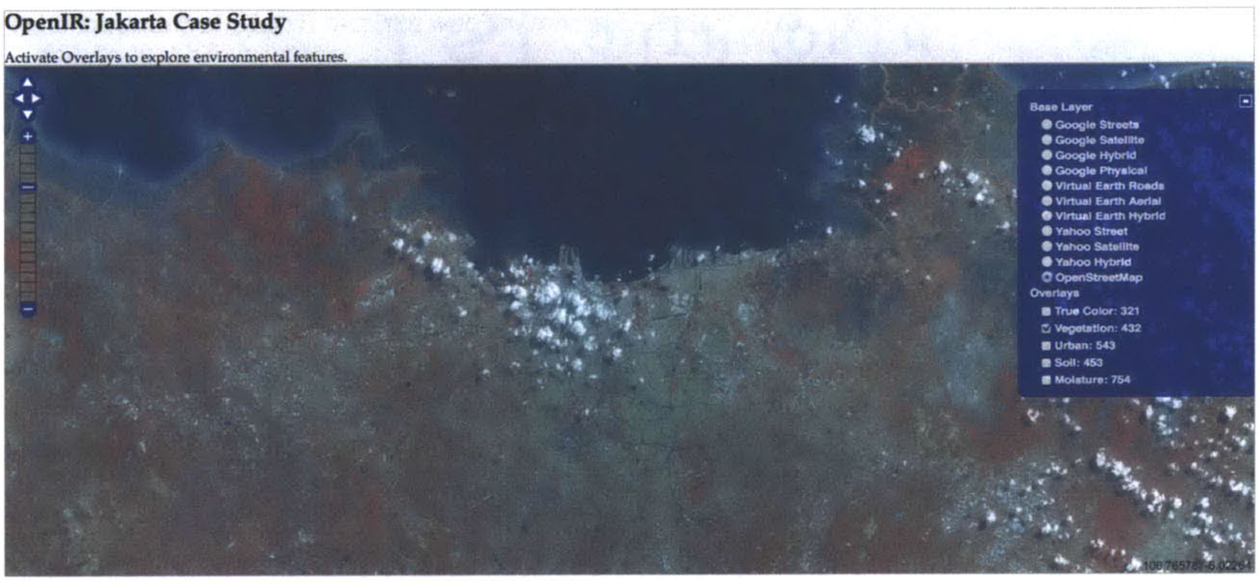
[26] University of Maryland Global Landcover Facility. http://glcf.umiacs.umd.edu

[27] "What are the band designations for Landsat satellites?" http://landsat.usgs.gov/band_designations_landsat_satellites.php

[28] Ned Horning, Center for Biodiversity Informatics, American Museum of Natural History: "Selecting the appropriate band combination for an RGB image using Landsat imagery." http://biodiversityinformatics.amnh.org/file.php?file_id=145

- Soil: Bands 5/4/3
- Moisture/Irrigation: Bands 7/5/4 (all non-visible bands)

I manually composited these band combinations for Jakarta and New York City using Photoshop's color channel tools, then saved each combination as a GeoTIFF file.[29] After that, we re-applied geographic metadata using the GeoTIFF/libgeotiff command line library, and then used GDAL to convert the GeoTIFF to a nested directory system of Tile Map Service[30] images, with an HTML file to call these images in a web map. Of particular use was the "GDAL2Tiles" script[31].



First deployment of OpenIR data in a web environment, January 2013.

iOS

After processing initial IR band combinations in January 2012, we experimented with building an iOS client for iPhones, iPods and iPads. We built this client on a sample software project for displaying OpenLayers TMS imagery.[32] Working with iOS was a rapid prototyping choice; we find it easier to build attractive software in a closed development environment like the iOS software development kit. However, even though smartphone, tablet, and mobile device usage is rapidly increasing throughout the world, we eventually decided to freeze this project in favor of a more widely deployable web approach.

---

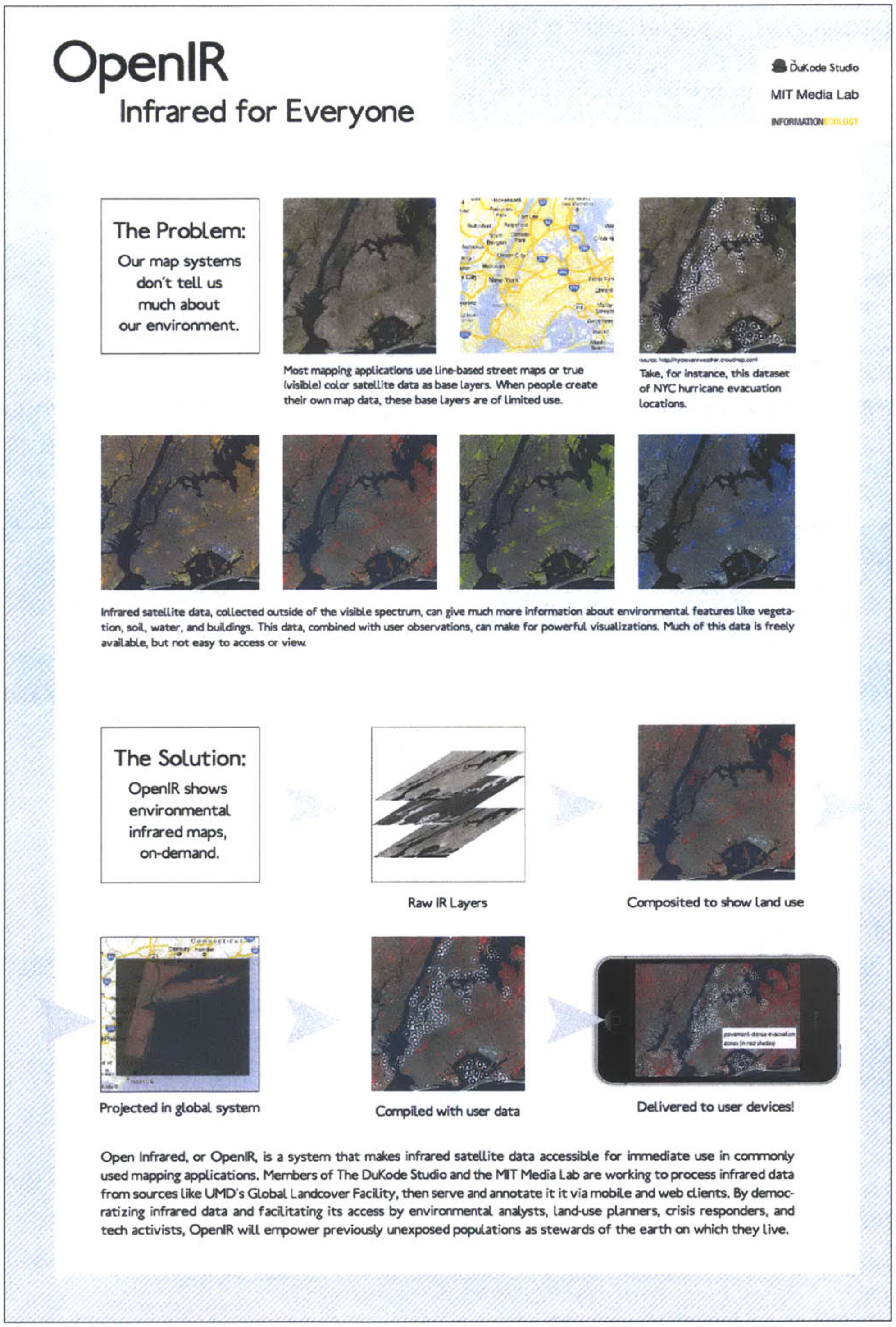
[29] GeoTIFF. http://trac.osgeo.org/geotiff/

[30] Tile Map Service Specification. http://wiki.osgeo.org/wiki/Tile_Map_Service_Specification

[31] GDAL2Tiles. http://www.klokan.cz/projects/gdal2tiles/

[32] Apple-WWDC10-TileMap. https://github.com/klokantech/Apple-WWDC10-TileMap

Nevertheless, this project was a useful way to start working with IR data in a software context. We also developed OpenIR's first concept poster, which shows the iPhone as the final delivery device.

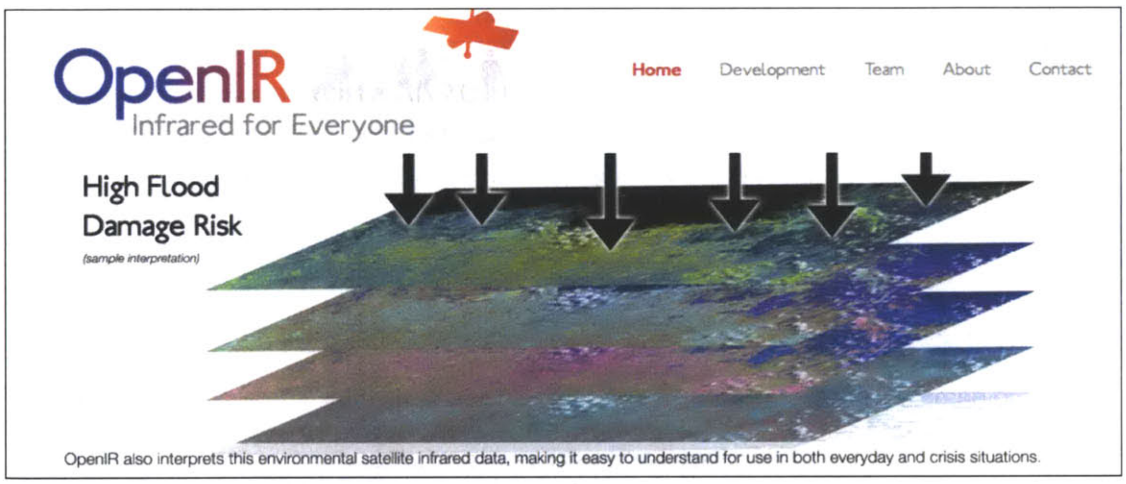


First poster for OpenIR, from January 2012. The poster explains IR maps, how they differ from other types of maps, and shows their display on mobile devices.

IDEAS Team Forming, Conceptualization

In February 2012, I started searching for team members for OpenIR to compete in the IDEAS challenge. At an IDEAS "generator dinner," I met Juhee Bae, an experienced community developer, mapper, and student in urban studies and civil engineering. She joined OpenIR as a volunteer, and later student Aziz Alghunaim joined as an undergraduate researcher in computer science. At this point, I had to teach the new team members about Infrared satellite data, its compilation, and its use, which was a useful process in considering how to build literacy in the general public.

My initial explanations to my new collaborators included a NASA Landsat Tutorial,[33] an interactive on the combination of infrared bands,[34] a demonstration of how IR satellite data is used to tell environmental stories,[35] and an explanation of commonly used Landsat band combinations.[36]



Screenshot from original http://openir.media.mit.edu web site, containing text, movies, and posters explaining the OpenIR concept.

For Aziz, my explanations also included training on OpenIR's existing code repositories, the software ImageMagick and GDAL, and the OpenLayers map library. At this time, we set up OpenIR on its own dedicated Media Lab virtual machine (http://openir.media.mit.edu) with Ubuntu Linux, and installed the basics for a web server: Apache[37], MySQL, and PHP. We

---

[33] Landsat Tutorial - NASA. http://landsat.gsfc.nasa.gov/education/tutorials.html

[34] Band Combination http://biodiversityinformatics.amnh.org/tool.php?content_id=141.

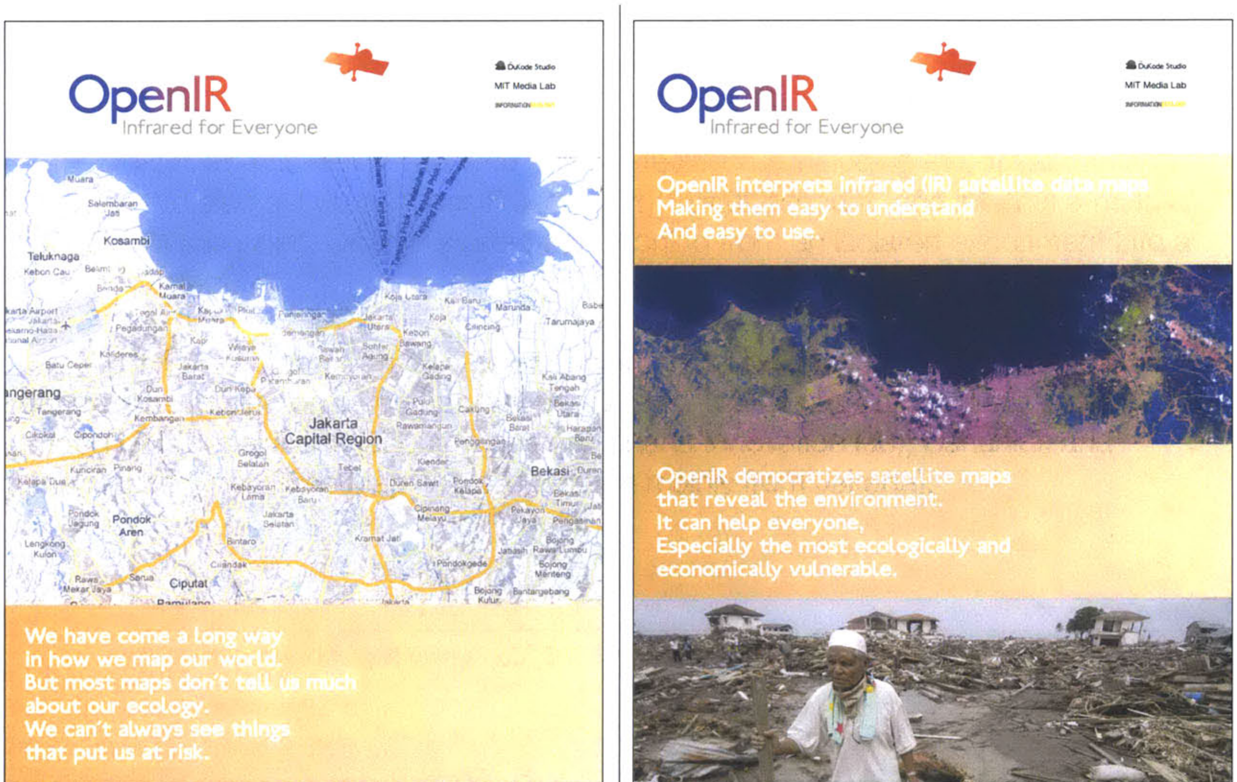
[35] Science Bulletins, Bio Visualization: Human Footprint. http://sciencebulletins.amnh.org/?sid=b.v.pearl_river_delta.20070926&src=l

[36] Landsat Band Cominations. http://web.pdx.edu/~emch/ip1/bandcombinations.html

[37] Apache LAMP server: https://help.ubuntu.com/10.04/serverguide/C/httpd.html.

set up GUI-based access software for these installations, including PHPmyadmin and ISPconfig. We also set up an OpenIR e-mail list, Twitter account, and internal e-mail forwarder. At this point, OpenIR started transitioning from an isolated experiment into a legitimate project.

Using the IDEAS competition as a helpful scaffold, Juhee and I started to ask related organizations to sign on as community partners for the OpenIR project. Initial key partners included UN Global Pulse (about to open a Jakarta innovation lab), Ushahidi, and Humanitarian OpenStreetMap Team (operating mainly from Jakarta).

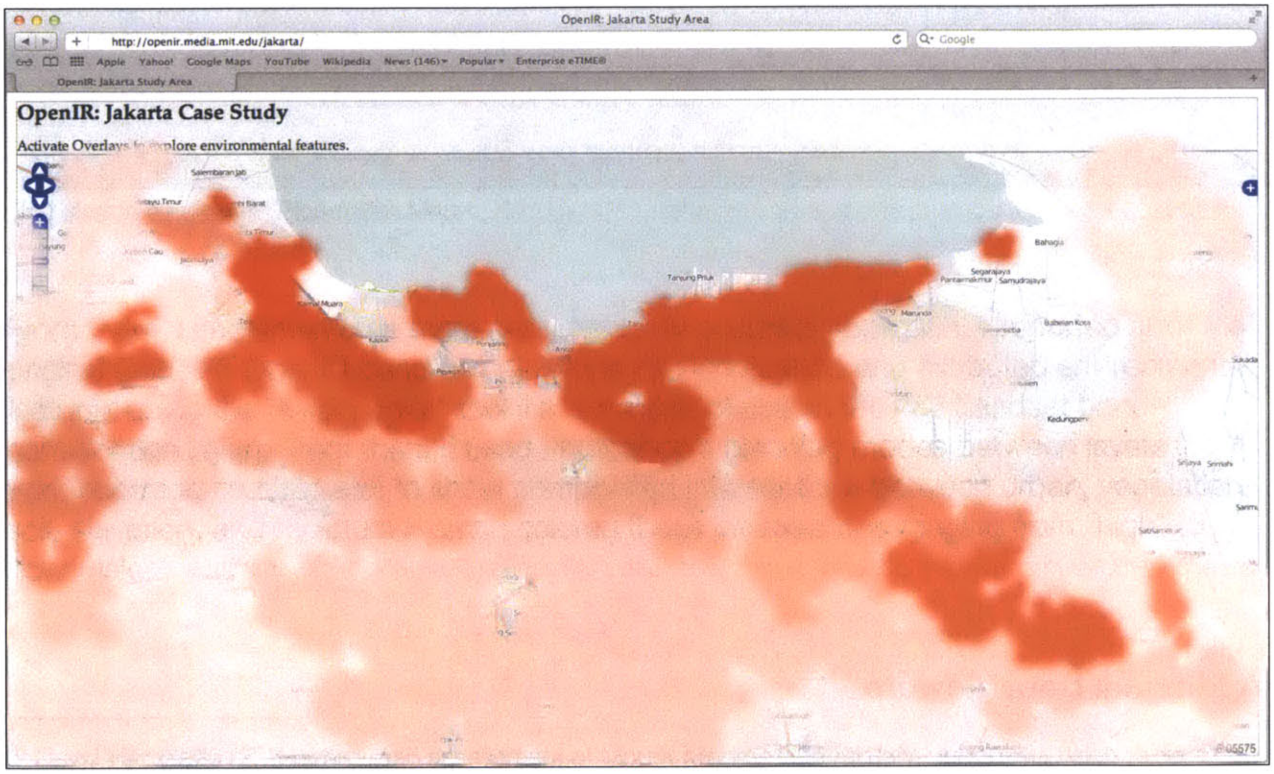


Panels from OpenIR concept posters, developed for IDEAS Global Challenge Judging session in April 2012. Much of February to May 2012 was spent connecting with community partners and refining the OpenIR concept.

First Risk Maps: Manual Construction

Environmental features are usually just one factor in evaluating disaster risk for a populace; social, architectural, municipal, and other factors also play a role. However, after reading about simple risk maps, compiled with IR and environmental data, to gauge volcanic risk on Indonesia's Mount Merapi,[38] it seemed that a simple, quickly constructed, accessible risk map could also play a useful role in emergency preparedness, especially in combination with other quickly constructed disaster maps like crowdsourced crisis maps.

After opening a screenshot of the Jakarta region in Adobe Photoshop, I quickly, manually blocked out areas that might be at higher risk (i.e. low-lying urban areas near water) and lower risk (i.e. high-lying vegetative areas far from water). This quick sketch is below.

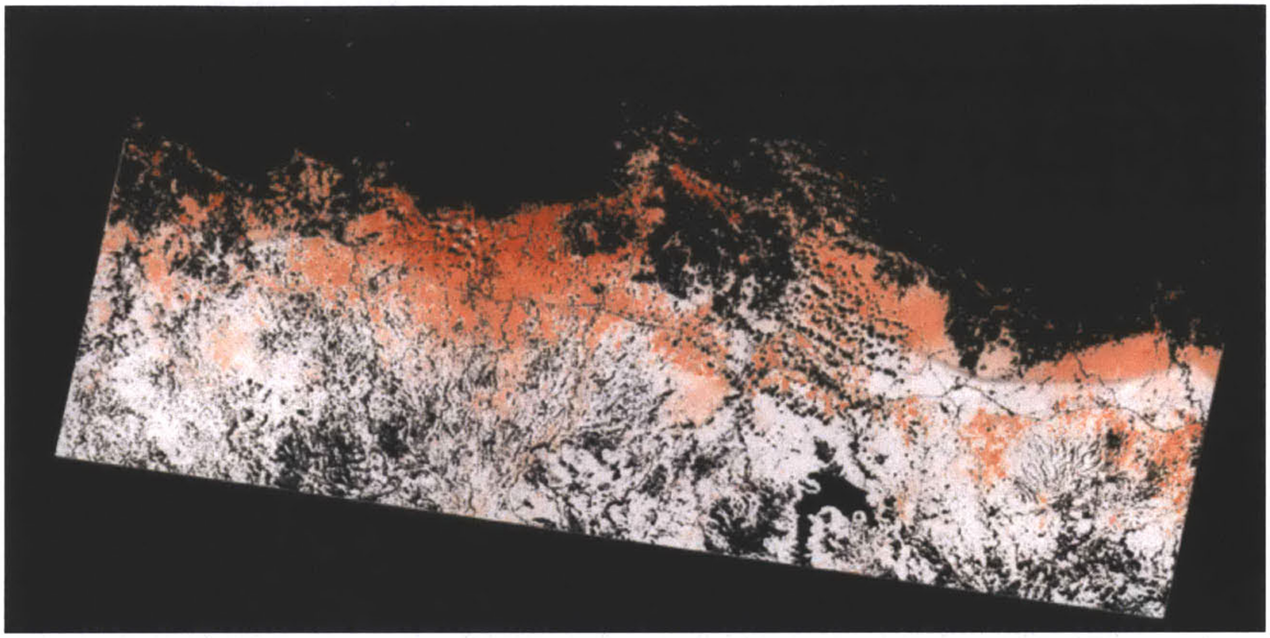


First sketch for Jakarta Flood Risk Map.

From there, I implemented a more geographically accurate approach. I imported all of the original GeoTIFF OpenIR band combinations into Photoshop and extracted environmental features using the "select color" tool (i.e. selection of pink in the 453 Landsat band combination layer). From there, I used Photoshop's blending modes between layers (i.e. add, subtract, multiply, etc) to show composited intersections between urban, vegetation, soil, elevation, and moisture layers. I colored these intersections ranging from "high" to "low" risk.

[38] David Harris. 2012. "Hazard map assessment of Mount Merapi, Central Java, Indonesia using remote sensing." http://earthquake-report.com/2012/01/11/hazard-map-assessment-of-mount-merapi-central-java-indonesia-using-remote-sensing/

This "Jakarta Flood Risk Map" manually compiled, was then converted to OpenLayers TMS images and added to the OpenIR Jakarta Data Viewer. A few months later, I would work with researchers to write out, refine, and automate the algorithm I used to manually compile this map (See Chapter 6).



First Jakarta Flood Risk Map, based on OpenIR inputs for urban surfaces, vegetation, soil, moisture, and elevation. Photoshop layers and blend modes were used to manually compile this map.

Additional Data Viewers

After the original OpenIR Data Viewers were compiled for Jakarta and New York, a few other data viewers were compiled during Implementation Phase 1. For the purpose of showing OpenIR in a local context, I made a Data Viewer for the Boston area.[39] At the request of environmental journalist Alister Doyle, our team also compiled a data viewer for London, his home,[40] and for Tuvalu, the lowest-lying land on planet Earth.[41] Implementation Phase 1 concluded with several IR Data Viewers, each showing the unique ways that human-built landscapes marked the natural environment.

---

[39] OpenIR Boston Data Viewer: http://openir.media.mit.edu/main/prototypes/Boston/.

[40] OpenIR London Data Viewer: http://openir.media.mit.edu/main/prototypes/London.

[41] OpenIR Tuvalu Data Viewer: http://openir.media.mit.edu/main/prototypes/Tuvalu.

Chapter 5: Front-End Evaluation Interviews
June 2012 to August 2012

The period between May 2012 to January 2013, bookended by the MIT IDEAS award and the Indonesia trip, was a time of intense development for OpenIR. At a "Winners Retreat" for the IDEAS award in late May, my collaborators Ilias Koen, Juhee Bae, Aziz Alghunaim, and I realized that we faced a lot of big questions and decisions, including:

- would it make more sense to process a large set of imagery for a small geographic area, or just one kind of image for a global dataset?
- should we focus more on data processing or user interface?
- which is more useful: on-demand processed data with a long wait time, or pre-processed data that is less recent?
- who has the most need for OpenIR?

My team and I shared a long-term vision for OpenIR, but we did not quite agree on short-term decisions and goals. We did unanimously agree that we needed a better picture of how scientists, mappers, funders, and practitioners perceive and use remote sensing products. To help with this, I wrote a set of Front-End Interview questions, mostly focused on usability. Undergraduate researcher Juhee Bae and I facilitated interviews with more than a dozen stakeholders, while Ilias Koen, Aziz Alghunaim, and visiting high school intern Faisal AlHumaidan listened and added additional questions.

In addition to the interviews summarized below, Kristin Cook-Turner of State Farm and Celina Agaton of UCNIID-Southeast Asia (University and Councils Network on Innovation for Inclusive Development in Southeast Asia) provided some useful guidance.

Usability Interview Summaries

<u>Kate Chapman, Acting Director, Humanitarian OpenStreetMap Team (HOT)</u>

Funded by the Indonesian and Australian governments, HOT's Indonesia pilot project is focused on expanding OpenStreetMap's accuracy in Jakarta and other parts of Indonesia, with a particular focus on mapping to mitigate the damage from flooding, tsunamis, and earthquakes. Kate Chapman has led this effort for the past two years, and has been based in Indonesia for the past year. Having taught many mapping workshops, she noted that "people can tend to be a little less excited about IR" than other kinds of maps.

However, she did point out that there are new audiences for environmental data amongst younger, web- and map-savvy Indonesians, many of whom are concerned with the destruction of their country's rainforests and orangutan habitats. "There are issues with logging and palm oil plantations, but we haven't been involved. There are just so many earthquakes, floods, and tsunamis."

When asked for advice on prioritizing local over global data processing, she said that local, detailed processing would be a better approach -- "A use case is better than a bunch of dispersed data." In her line of work, which often deals with emergency management and flood/contingency planning, she said that spatial resolution is most important, followed by temporal resolution. Spectral resolution is useful but for her, the least important.

Concerning user populations, Kate Chapman said that local NGOs and advocacy groups would make most use of OpenIR. Indigenous populations and low-lying populations (those at low sea level) would be most helped by OpenIR, she said.

Robert Kirkpatrick, Director, UN Global Pulse

HOT and UN Global Pulse were OpenIR's first community partners, since both organizations have pilot projects in Indonesia. Robert Kirkpatrick and some of his team members met with OpenIR in June 2012. He said that local maps are most important to UN Global Pulse, because the organization is taking a local approach through their "Pulse Labs," initially in Jakarta and Kampala.

He said that human input to maps is important, and he was particularly interested in forming hypotheses using passive data, i.e. data that people accumulate through their everyday device use, without active submission to a data collection mechanism.

When asked about who would most use IR map software, he answered "governments and those who help them prepare for vulnerabilities, focus on the local communities." He emphasized the importance of getting "the information to the local communities because they have never seen it before."

He also advised the OpenIR team to "definitely work on a mobile software especially for Jakarta" because so many people there use smartphones.

Lela Prashad, Chief Data Scientist, NiJel.org

Lela Prashad is a technical adviser to OpenIR, having worked with remote sensing data through Arizona State University and NASA Jet Propulsion Laboratory. During the UN

Global Pulse + OpenIR meeting, she suggested that for global datasets, "you'd want to 'can' the data, process it as it appears, because nobody wants to wait."

Nathaniel Manning, Director of Business Strategy, Ushahidi

Many members of Ushahidi have provided helpful information to OpenIR, including US-based Nathaniel Manning. He said that detailed local data is probably better than a generalized global dataset, as it provides more accessibility to local residents.

For crisis applications like those of many Ushahidi deployments, he said that temporal resolution is most important. Spectral resolution is relevant for applications concerned with several environmental features at once, like disease tracking in crops.

In terms of data processing, he said that the easier and more streamlined data availability is for the user, the better for the actual crisis situation, especially for users/deployers in the midst of crisis. On the other hand, if recent data is sacrificed at the expense of user-friendliness, on-the-ground time and resources can be wasted.

Governments, NGOs, crisis responders and social entrepreneurs are most likely to use OpenIR, he said. When we showed surprise at his mention of this last group, he said, "look at what social entrepreneurs have done with GPS data and weather data!"

Heather Leson, Director of Community Engagement, Ushahidi

Also from Ushahidi, the Toronto-based Heather Leson is one of the most engaged and accessible members of the core Ushahidi team. In addition to contributing a Front-End interview, she connected OpenIR to Willow Brugh of Geeks Without Bounds and a series of "hackathons" that were quite useful for OpenIR's autumn 2012 implementation goals (see Chapter 8).

She shared a particular interest in adding OpenIR to the "sensemaking" of crowd sourced map data. She had an ultimate vision, like many people working with interactive map data, of a deeply layered, multi-sourced map containing citizen data, sensor data, and satellite data, all available for analysis and cross-analysis. She pointed out that while Ushahidi is rooted in election monitoring, many of its members are keen to see how scientists and "environmental people, etc" can use environmental data with crowd-map software.

She added: NGOs and governments change slowly. There's probably a 3 year time lag of adoption with NGOs and governments, so the most flexible and most likely to adopt OpenIR are crisis responders and journalists.

Sumeeta Srinivasan, Perceptor, Harvard Center for Geographic Analysis

Dr. Srinivasan teaches a number of geographic systems and statistics classes at Harvard. She is familiar with the geographic work at MIT, where she received her Ph.D. Her work is conducted at regional or larger scales. It is not conducted with social or crowd-sourced data, but it does make extensive use of many remote sensing products.

On the topic of accessibility, she said that while working with the World Bank, which has a big group doing risk reduction, she "was trying to persuade them that they didn't need to use proprietary risk analysis software."

For her own work, "spectral resolution is important" to "do classification. I use impervious surface band combinations, health of vegetation and agriculture, especially some of these cities in India and China have a huge water problem."

Finally, she said that insurance companies may most use OpenIR, although "a lot of these government planning agencies need it but don't realize it."

Kirk Morris, Standby Task Force

The Standby Task Force (SBTF), a loose, international volunteer group that deploys crisis- and crowd-mapping technologies when "activated" by municipal organizations, has several members who have provided guidance to the OpenIR project. Kirk Morris, an indefatigable US-based crowd mapper, is one of SBTF's most active participants and was one of its leaders.

He emphasized that SBTF only responds to highly organized groups who need specific information from a crisis map. Few of these responses take place in the United States because "every state has its own [crisis] entity, its own emergency response protocols, even with FEMA at the top. So who do you talk to...? Who are you serving? Hard to say in the U.S." Instead SBTF works with international groups like UN OCHA, the World Health Organization, and Amnesty International.

He told the story of how "for the first 4 days of the SBTF Libya Crisis Map [in which the Libyan government was overthrown], I was on for 48 hrs continuously." In this time, he said that "no one could answer how they [Libya] went from a few refugees to 4000." Kirk Morris's avid Twitter use allowed him to see that "certain checkpoints were choked along the coast, refugees couldn't get to Tunisia." If he could have identified this scenario through IR data instead, "that would be hugely valuable."

Regarding the OpenIR's question of pre-processed versus more recent, rougher data, he said that for SBTF, processing power is not an issue since volunteers are remotely located and often have high-speed Internet connections. However, for "folks on the ground, rapidity of download is an issue." In SBTF, its Analysis Team helps bridge this gap, by narrowing and refining large data streams to what is most relevant for the crisis.

Science Interview Summaries

In addition to the usability-focused Front-End Interviews, we also conducted longer, more in-depth sessions with several scientists using remote sensing with crowd and crisis applications. These sessions were less uniform than the usability interviews, since we wrote a unique list of tailored interview questions for each scientist.

Einar Bjorgo, Senior Specialist, UNOSAT

During a brief conversation with FEMA's Ted Okada, he suggested that we take a look at the Ph.D. work of Dr. Bjorgo, who is a senior scientist at UNOSAT, the section of the United Nations that delivers satellite imagery and analysis for crisis and relief applications.

We asked Dr. Bjorgo about UNOSAT's data sources and opinions on commercial versus free satellite data. He answered that some of the best support UNOSAT has given is based on daily MODIS (Moderate Resolution Imaging Spectroradiometer) data. For instance, MODIS was used in UNOSAT's response to the 2010 Pakistan Floods. On the other hand, very detailed information must come from commercial providers. UNOSAT has a small budget for commercial images, and prices of commercial imagery have dropped enormously in recent years. UNOSAT also has access to data through International Space and Major Disasters, an international consortia that freely exchanges satellite imagery during major disasters. This consortia bridges the public and private satellite worlds.

Regarding the automation of risk map derivation, he said that this is possible to a certain extent. MODIS data is used to automate fire detection. For flood detection, UNOSAT employs a semi-automatic process. Most risk detection processes must be fine-tuned according to the environmental and architectural context. In terms of more complex analysis (change detection, damage detection, etc), the human eye is still the best analysis tool. Some procedures can be semi-automated, but there is always a limit.

As we finished our interview, Einar Bjorgo commented that UNOSAT is focusing on internally displaced people in Somalia (IDP). There are instances of fraud, in which aid money is awarded to groups who set up dozens of empty tents to give the illusion of large displacement camps. If IR (thermal and vegetation signatures) data was applied to this situation, it could help enormously to to indicate where people are actually living.

Ken Duda, Senior Scientist, NASA/USGS

Our adviser Lela Prashad suggested that we contact Ken Duda, a scientist applying IR data from the ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) satellite instrument to US crises like Hurricane Katrina. He is also involved with United States Geological Survey Hazards Data Distribution System (USGS HDDS), a database of recent hazard imagery.

His focus on ASTER stems from a long experience with it, even prior to its 1999 launch. He named ASTER's advantages, including "reasonable" resolution (up to15 meter pixel resolution) and multiple bands (for instance, short-wave infrared for burn scar and mineral mapping). On of ASTER's disadvantages is that data is not acquired at a high temporal rate. ASTER is "good for a bit of a finer look—roads, buildings, but not people." Landsat has similar capabilities, with some overlap in the bands.

On the topic of OpenIR's risk mapping, Ken Duda commented that automated derivation of hazard maps may be of value in a real time situation, as long as they are made very rapidly.

Ned Horning, American Museum of Natural History

When I worked at the American Museum of Natural History (AMNH) from 2005 to 2011, I was slightly acquainted with Ned Horning, the Director of Applied Biodiversity Informatics at AMNH's Center for Biodiversity Conservation. I also became quite familiar with his clear, useful interactive tutorials on Remote Sensing.[42] After seeing several of his posts on the email list for the Public Laboratory for Open Science and Technology, a group with close ties to the Media Lab, I set up an interview with Ned Horning to gauge his experience with remote sensing, citizen science, and their environmental applications.

He mostly uses Landsat data for his environmental studies, because it has "a nice resolution. Having those extra IR bands are quite useful, especially for cloud cover detection." His studies are regional or continental, so he doesn't need high spatial resolution.

He adds that for predictive modeling, Landsat can be extremely useful, because it offers such high spectral resolution. "You want to know the vulnerability layer for a large province or country, so Landsat is good. For disaster recovery, Landsat is quite nice for that. For response, not so much."

---

[42] AMNH CBC Remote Sensing Tutorials, Interactive Tools. http://biodiversityinformatics.amnh.org/index.php?sectionnav=&globalnav=&section_id=10.

He is an avid user of Random Forests, an algorithm that involves training a software agent to recognize environmental features based on satellite imaging. He likes Random Forests because "it's very easy to run, it's forgiving to outliers, and it's easy to parameterize. Other algorithms require a lot of special testing and setup."

We showed him our risk index algorithm [see previous chapter] and asked if he thought Random Forests could be a good fit for our plan to develop the risk algorithm. He answered that Random Forest requires active training from the user to the software agent, so for an automated process like OpenIR's risk algorithm, it would not make sense to apply Random Forests.

<u>Michael Flaxman, MIT & GeoAdaptive, LLC</u>

When I first arrived at MIT, I approached Dr. Flaxman, then a professor in the MIT Department of Urban Studies and Planning, for suggestions on GIS programming courses. One year later, I approached him with the OpenIR team to solicit his thoughts related to our project. He offered some of the most thoughtful, incisive commentary, from years of working in computational and participatory planning for MIT and for ESRI, one of the largest GIS software companies. He now runs a small consultancy in San Francisco called GeoAdaptive, LLC.

He explained how important citizen participation and vulnerability analysis is to his urban planning practice. "It's always been challenging in developing economies," he said, "because there's less spatial data infrastructure and less access to open data over all." But this is changing, and he said that OpenIR could be a big piece of that. "The things you still can't get from remote sensing is an issue, but crowd sourcing can help."

Based on his experience as an architect of the ESRI Model Builder software, he said that OpenIR's approach to data processing could be improved: "it would be a useful interface for a data dependency diagram to be used for building the model.
The structure of most of the vulnerability models they use are in the realm of 20-50 nodes at most. They want to be able to repeatedly revise and bring audiences different parts of the diagram. At some point, it becomes too complicated to directly relate the bands to the classifications and other abstractions."

Regarding OpenIR's focus on Indonesia, he said that "most areas of the developing world are developing rapidly. There's a need to develop in the places with less risk and more infrastructure." He ascribes to the idea of using risk analysis to drive infrastructure, and settlement will derive from there.

He strongly liked OpenIR's idea of a generic set of global risk maps, "but what's tricky is that these risk maps need to be customized for a specific region. That customization phase (weighting/rating) is the hard part. There are cultural and built overlays to these constructs. You want to be able to figure out where 20% of people are at highest risk. If you express that at relative rank, this is most useful. You want at least 5 degrees of risk in a local region: where is bad, where is good, where are people most at risk."

But there needs to be a participatory element to vulnerability mapping, he said. "A lot of parts of the developing world are centrally planned, people wait to be handed information from the top." A big part of his consultancy work is to encourage people to think differently about planning. "One of the great things about maps is that a high school student can develop an alternate proposal and send it up the chain," thus flipping the paradigm of information flow. A primary objective with the kinds of tools he builds and uses is to empower people to "create their own version of a map, illustrate their own vulnerabilities, and share it." He says that in OpenIR, the same tool that can help tweak classification can also help in the weighting/rating of vulnerability assessment. "This would be much more participatory."

Takeaways

These interviews helped our team make the decision to focus on a local geographic area with multiple types of data, instead of processing a global area with one type of data. This supports our plan to focus on local communities in Indonesia and on the U.S. east coast. We learned that key players share our vision of interactive maps that more richly integrate social and satellite imagery for deeper sense-making. We also learned that the dataset we are focused on, Landsat, is best for recovery instead of rapid-response.

In regards to mapping risk, these interviews helped us to better understand the classification context in which environmental risk is assessed. Ned Horning discussed training versus automated algorithms and pointed out that algorithms like Random Forests do not apply to our approach. Michael Flaxman also took a look at our risk assessment approach and pointed out that while it is valid, it does not allow for a complex set of inputs, nor does it allow for the participatory modification that is so important in community planning.

Many issues were raised, some feasible, some not feasible to address within the scope of this thesis. After these interviews, it was clear that while our risk map algorithm is simple, it was taken seriously by expert practitioners who made helpful suggestions for improving and contextualizing the approach.

Chapter 6: Implementation Phase 2
June 2012 to August 2012

Having manually built a simple prototype Data Viewer to demonstrate in spring 2012, the summer 2012 months were spent developing processes to automate Data Viewer construction. This development included

- converting the basis of the existing data viewer from the OpenLayers to the Leaflet map library
- evaluating software from Arizona State University (ASU), at the suggestion of OpenIR adviser Lela Prashad
- diagramming and refining the Risk Index mapping algorithm used to manually construct the spring 2012 risk map
- converting this algorithm into a Python/Bash script
- writing Python/Bash scripts to collect, correct, convert, and display IR satellite data in a web browser.

By the end of the summer, undergraduate researcher Aziz Alghunaim and I had streamlined our prototype so that it looked better, ran more smoothly, offered an auto-generated Risk Index map, and could be replicated for other geographic areas.

Background: Data Acquisition for Laypeople

Navigating through and processing raster map imagery can be a significant challenge: the file sizes of high spatial resolution raster imagery are quite large, especially in comparison to point data (i.e. geotweets) or line data (i.e. political boundaries).
Combining raster maps can be a slow process if one or both maps must be reprojected into matching map systems. For laypeople only looking for high spatial resolution in raster imagery, the use of a web map system like Google Maps, Bing, or Yahoo can be enough.

However, most raster maps are photographs of some kind, making them useful for examining a landscape's change over time. But again, the large file sizes of high spatial resolution imagery can grossly complicate the process of introducing temporal resolution. Combined with issues of privacy (i.e., do I really want a daily 1-meter resolution image of my apartment's patio available to the public?), it is no wonder that few free map systems integrate high spatial and temporal resolution together.

On the other hand, spectral resolution affords the layperson the opportunity to see environmental features more clearly delineated. At coarser spatial resolutions, privacy issues can be avoided and environmental analysis can be supported. But laypeople seeking imagery outside of the visible spectrum not only face the file size challenges of all raster imagery, they also also face challenges of selecting and compositing spectral bands into "false color"[43] images: the infrared spectrum is much wider than that of the visible spectrum, and specialized tools are needed to composite spectral bands. There are also interoperability challenges. Often, spectral imagery is available only in formats like HDR (High Dynamic Range), which are not viewable with most common graphics tools.

With these challenges in mind, we focused much of this implementation period on developing a processing pipeline that would give the most access to datasets with high spectral, as well as spatial and temporal resolution. We evaluated several APIs for making this data available, such as the Davinci and J-Earth systems, and we also evaluated several kinds of IR datasets and their APIs as available.

Evaluation of ASU's Davinci

Davinci is a software package, developed at Arizona State University, to manipulate and view spectral data. We evaluated this software to gauge whether Davinci could replace and/or improve with tasks for which we were using the command line packages GDAL (Geospatial Data Abstraction Library) and ImageMagick. In general, we use GDAL for all geographic processing (reprojections, conversions, tiling, etc) and ImageMagick for all image processing (i.e. color channel assignments, composting, etc).

Our evaluation plan was to load a satellite image into Davinci and test how the image could be manipulated. Since imagery from the ASTER satellite instrument is commonly used with Davinci, we downloaded the relevant library[44] and loaded it:

```
dv>> source("pathto/davinci_library/msff.dvrc")
```

From there, we started to work with the library's load_aster() function, but encountered a few prohibitive errors:

- The error
  Failed HDF data extraction. Exiting...

---

[43] "NASA Starchild Question of the Month: What is meant by false color imagery?" http://starchild.gsfc.nasa.gov/docs/StarChild/questions/question20.html

[44] Davinci Library: http://davinci.asu.edu/index.php?title=Download_Davinci#Davinci_Library

was, as a member of the Davinci team from ASU suggested, solved by change the $DV_GDAL environment variable to point at the path with the gdal programs.
The command you would use for that is:
putenv("DV_GDAL","/Library/Frameworks/GDAL.framework/Versions/Current/Programs/")

- The error
  Illegal range value.
  Illegal range value.
  length: Variable not found: names
  error: Variable not found: lines
  struct, 3 elements
  anc: struct, 1 elements
  path: Text Buffer with 0 lines of text
  bandlist: Text Buffer with 0 lines of text
  data: struct, 0 elements

  Had to do with reading the HDF file format. There is a Perl script that load_aster calls (filename: "aster4davinci.pl"). It must be obtained and copied to the right directory
  sudo cp pathto/aster4davinci.pl /Applications/davinci.app/Contents/Resources/library/script_files

- There are also errors if GDAL is not configured to support HDF format, which is how many ASTER images are stored.

After fixing the errors above, the first error occurred again. The developers at ASU provided some very helpful assistance, but this evaluation was proving to be too effort intensive with few useful results, so we decided to shelf the use of Davinci.

Evaluation of ASU's J-Earth

J-Earth is a Java-based desktop program that gives user access to a wide range of spectral map datasets. In some cases, data is acquired on-demand from various sources and servers.

For OpenIR's purposes, I wanted to gauge whether J-Earth, or some of its algorithms, could be used to better automate data acquisition, particularly from the ASTER satellite instrument. Unfortunately, after evaluating the software, talking with the team, and evaluating how ASTER is acquired from government databases like NASA Reverb, automated acquisition is not an option, especially since control of the satellite instrument is highly manual. All ASTER image requests must be made manually.

Data Beyond Landsat

For initial OpenIR prototypes, imagery from the Landsat instrument was chosen for its spectral resolution, moderate spatial resolution, clarity, and ease of acquisition. However, we evaluated data beyond the Landsat source, including ASTER, MODIS, and IKONOS data. OpenIR advisor Lela Prashad, Katie Baynes of NASA Echo, and Ken Duda and Jan Wilson of USGS LPDAAC, helped with this evaluation.

ASTER (via NASA ECHO Reverb)
ASTER offers higher spatial resolution imagery than Landsat. I tried "ordering" a large amount of ASTER data through the NASA Reverb ECHO data acquisition system. After a system error, LPDAAC, the central storage facility for ASTER data, rejected all my orders. After speaking with Jan Wilson of LPDAAC, I was told that it is not possible to automate our downloads of this data, which is available only via manual orders. With this, I de-prioritized the use of ASTER data in OpenIR, especially since the ASTER SWIR (shortwave infrared, approximately analogous to bands 5, & 7 in Landsat) sensor was disabled in 2008.

IKONOS (commercial data)
IKONOS is very high resolution data, with a spatial resolution of up to 1 meter per pixel. The European Space Agency no longer distributes IKONOS archive data. The archive is available through a paper-based order to E-GEOS in Italy. I wrote to the primary source of IKONOS, the private company GeoEye. Organizations can use GeoEye sample data for free and/or apply for an imagery grant for up to 500 sq. km of imagery. But it's unclear how often OpenIR would need to re-submit grant applications. It's also unclear if we can obtain much IR data: INKONOS only have bands in the visible and near-infrared spectra. Finally, while it's possible to acquire a few images for free, the only way that programs like OpenIR can automate IKONOS acquisition is to pay a large subscription fee.

MODIS
MODIS imagery has a lower spatial resolution than Landsat (about 60 meters per pixel), but it is collected much more frequently than any other publicly available satellite data image. While meeting at the UN Global Pulse office in June 2012, Lela Prashad suggested not to write MODIS off; this data is used for large-region monitoring. Considering that we already have scripts set up to access GLCF (UMD's Global Landcover Facility), it should be straightforward to incorporate the MODIS vegetation data that's on GLCF.

Automating Landsat Acquisition and Processing

As of the writing of this thesis, all OpenIR data viewers are compiled with data from the Landsat instrument. As mentioned previously, other datasets are useful but less ideal for the OpenIR prototypes. In June 2012, Aziz and I developed several scripts to acquire Landsat data and process it into a web-based map browser. The specific steps are:

- downloading Landsat IR bands from the [University of Maryland's Global Landcover Facility (GLCF)

- building band composites that highlight discrete environmental features

- color correcting and georeferencing the composites

- slicing & dicing the images up for caching

- building an HTML/Javascript viewer for the composites

These scripts are stored in the DuKode repository at GitHub.[45] The script "compositetile.sh." is the master script that will run the other scripts in the directory, and the "gdal2tiles" script is based on the Klokan-developed script included with GDAL.

In keeping with Landsat WRF (Worldwide Reference System[46]), the current location input method uses path/row, not city name nor latitude/longitude coordinates.

Algorithm for Making a Simple Flood Risk Index

To help automate the process of outputting a Risk Index based on the manually generated Spring 2012 Risk Map (see Chapter 4), I diagrammed and wrote out the algorithm for the manual compilation that drew from Harris's Indonesia-based hazard assessment[47].

This algorithm also draws on some key concepts of InaSAFE's software, namely that of *hazard, exposure,* and *impact.* For InSAFE, all questions are formulated in the form:

> *In the event of* **[hazard]** *how many* **[exposure]** *might* **[impact].**
> *For example: "In the event of a* ***flood*** *how many* ***buildings*** *might be* ***closed****?"*

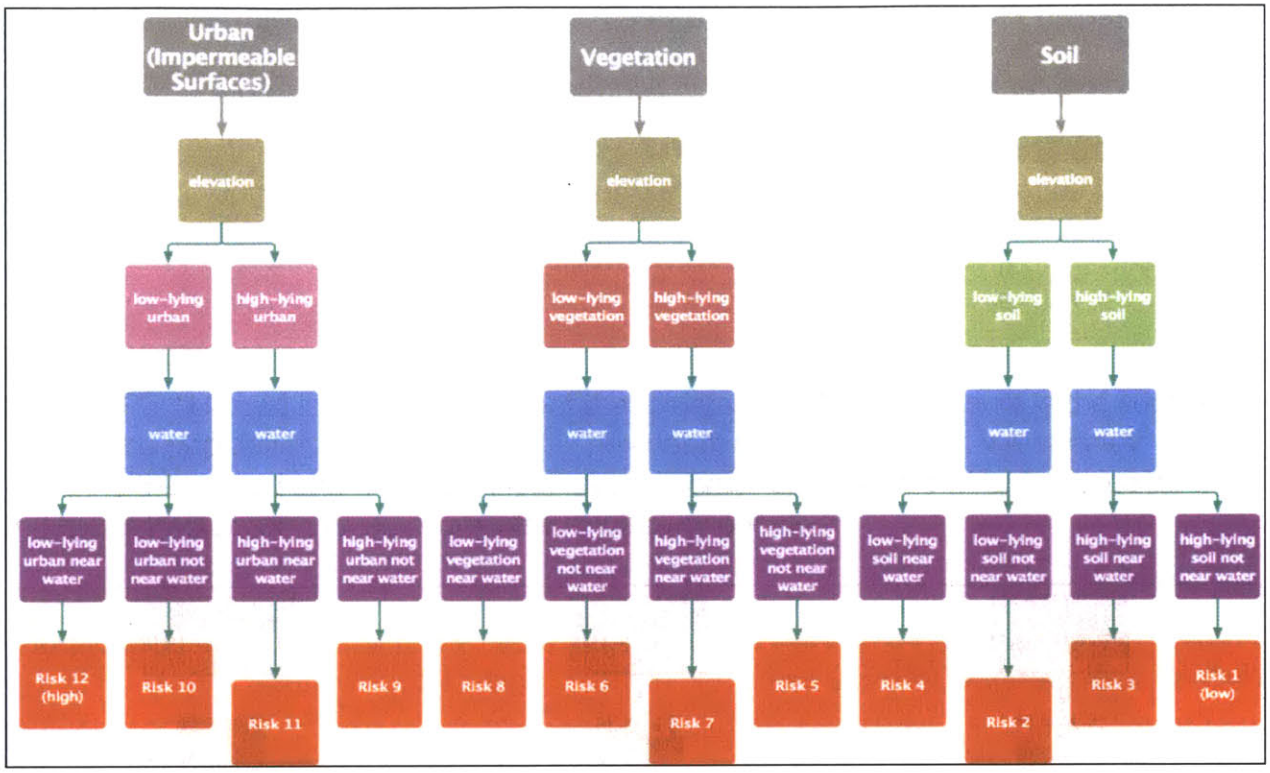
---

[45] OpenIR_ImageProcAndPrep: https://github.com/DuKode/OpenIR_ImageProcAndPrep

[46] Landsat WRF: http://landsat.gsfc.nasa.gov/about/wrs.html

[47] David Harris. 2012. "Hazard map assessment of Mount Merapi, Central Java, Indonesia using remote sensing." http://earthquake-report.com/2012/01/11/hazard-map-assessment-of-mount-merapi-central-java-indonesia-using-remote-sensing/

In InaSAFE, "Hazard is the physical event that creates the risk... Exposure is the sum of assets and population that are at risks... The impact ("might") function will spatially combine the hazard and exposure input layers in order to postulate what the impacts of the hazard will be on the exposure infrastructure or people."[48]



OpenIR Flood Risk Index, directed graph.

OpenIR applies the notion of spatially combining hazard and exposure to display potential impact. Like with InaSAFE, the hazard is a flood, and the algorithm is too simple to be called a "model."[49] But while InaSAFE calculates flood impact by spatially combining a *hazard* of historic flood data (i.e. 2007 flood extents) with an *exposure* of Jakarta building footprints, OpenIR calculates flood impact by spatially combining a *hazard* of low-elevation land with an *exposure* of impervious surfaces (most building materials are impervious surfaces). In using elevation as the basis of for *hazard*, OpenIR takes a similar approach to Hazus, which uses elevation imagery as a predominant factor in calculating flood hazard

---

[48] InaSAFE Tutorial — InaSAFE 1.9.1-final documentation. http://inasafe.org/tutorial-docs/tutorial.html

[49] The InaSAFE site specifically states: "It is important to note that InaSAFE is not a hazard modeling tool. Information on hazards needs to be provided either by technical experts, often from Government agencies, universities or technical consultants, or from communities themselves based on their previous experiences." http://inasafe.org/

via streamflow and overflow.[50] While OpenIR does not calculate streamflow as finely as Hazus, OpenIR does similarly calculate drainage by expanding the area covered by water. OpenIR makes a more conservative calculation, only expanding by 300 meters, while Hazus expands by about 1600 meters (1 mile). Hazus's calculation processes becom significantly more complex from that point, incorporating a wide array of hydrologic and socioeconomic data.

While OpenIR's approach to risk assessment is significantly simpler than Hazus's or even InaSAFE's, its uses data that is available for all landmasses on the globe, namely that of IR and elevation imagery. This makes the OpenIR approach a potentially useful starting point for most economically developing regions, which may not have the same level of municipal, infrastructural, and socioeconomic data as that in Hazus or InaSAFE.

As discussed in Chapter 5, we shared our algorithm with scientists, receiving preliminary validation and relevant, useful feedback, particularly from Ned Horning and Michael Flaxman.

<u>OpenIR Risk Index Algorithm</u>

Step A. Isolate environmental features. In the case of the OpenIR data viewer:
water (blue in Landsat 754),
urban (pink/magenta in Landsat 543),
soil (orange/brown in Landsat 453),
vegetation (red in Landsat 432).
elevation (white in Shuttle Radar Topography Mission elevation map)

<u>To process Risk Level 10 down to Risk Level 6:</u> the basic rubric is to cross-check urban areas (impermeable surfaces), which one can assume will be areas with the most humans and the most buildings, with areas of water and elevation. Areas with or near a lot of water are more prone to flooding, and areas of higher elevation are more protected from flooding.

Step B. Isolate Low Lying Urban Areas:
Elevation (on top), "lighten" with urban.
Delete everything but pink/magenta.

---

50 Silvana Croope, 2009. "Working with HAZUS-MH: A working paper submitted to the University of Delaware University Transportation Center (UD-UTC),"pp. 21-27. http://www.ce.udel.edu/UTC/Final-Working%20Paper-HAZUS-091028_rev.pdf

Step C: Isolate Higher-Lying Urban Areas:
Elevation (on top), “darken” with urban.
Delete everything but pink/magenta.

Step D: Isolate Low-Lying Urban Areas Near Bodies of Water.
Select Bodies of Water (darkest blue/black in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Urban Areas (from Step B) in this selection, delete everything else.
RESULT: Areas of Level 10 Flood Risk

Step E: Isolate Higher-Lying Urban Areas Near Sea.
Select Bodies of Water (darkest blue/black in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Urban Areas (from Step C) in this selection, delete everything else.
RESULT: Areas of Level 9 Flood Risk

Step F: Isolate Low-Lying Urban Areas Near Irrigated Land.
Select Irrigated Land (navy in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Urban Areas (from Step B) in this selection, delete everything else.
RESULT: Areas of Level 8 Flood Risk

Step G: Isolate Higher-Lying Urban Areas Near Irrigated Land.
Select Irrigated Land (navy in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Urban Areas (from Step C) in this selection, delete everything else.
RESULT: Areas of Level 7 Flood Risk

Step H: All Other Urban.
Any other Urban areas not isolated in Steps D-G.
RESULT: Areas of Level 6 Flood Risk

<u>To process Risk Level 5 down to Risk Level 1:</u> the basic rubric is to cross-check vegetation areas, which one can assume will be areas with fewer humans and buildings, with areas of water and elevation. We’re basically going through the same steps above, but with vegetation instead of urban (impermeable surfaces).

Step I. Isolate Low Lying Vegetation Areas:
Elevation (on top), “lighten” with vegetation.
Delete everything but red.

Step J: Isolate Higher-Lying Vegetation Areas:
Elevation (on top), "darken" with vegetation.
Delete everything but red.

Step K: Isolate Low-Lying Vegetation Areas Near Bodies of Water.
Select Bodies of Water (darkest blue/black in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Vegetation Areas (from Step I) in this selection, delete everything else.
RESULT: Areas of Level 5 Flood Risk

Step L: Isolate Higher-Lying Vegetation Areas Near Sea.
Select Bodies of Water (darkest blue/black in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Vegetation Areas (from Step J) in this selection, delete everything else.
RESULT: Areas of Level 4 Flood Risk

Step M: Isolate Low-Lying Vegetation Areas Near Irrigated Land.
Select Irrigated Land (navy in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Urban Areas (from Step I) in this selection, delete everything else.
RESULT: Areas of Level 3 Flood Risk

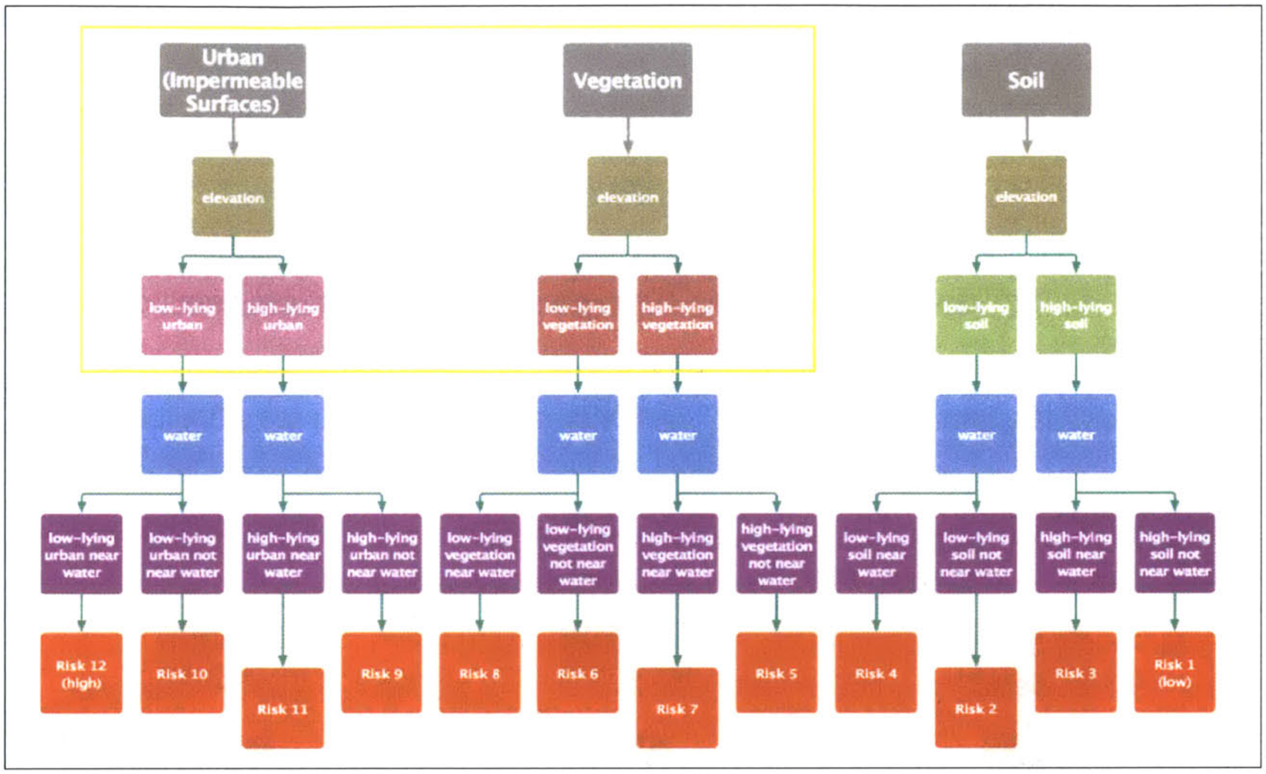
Step N: Isolate Higher-Lying Vegetation Areas Near Irrigated Land.
Select Irrigated Land (navy in Landsat 754).
Expand for drainage (10 pixels = 300 meters)
Keep all Low-Lying Vegetation Areas (from Step J) in this selection, delete everything else.
RESULT: Areas of Level 2 Flood Risk

Step O: All Other Vegetation.
Any other Vegetation areas not isolated in Steps K-N.
RESULT: Areas of Level 1 Flood Risk

Step P: Recolor Areas of Flood Risk (Level 1-10) with a white-to-red gradient.

Auto-Generated Risk Map with Urban, Vegetation, and Elevation Inputs

The manually-generated risk map showed 12 risk levels, as explained in the algorithm above, and in the decision tree below:



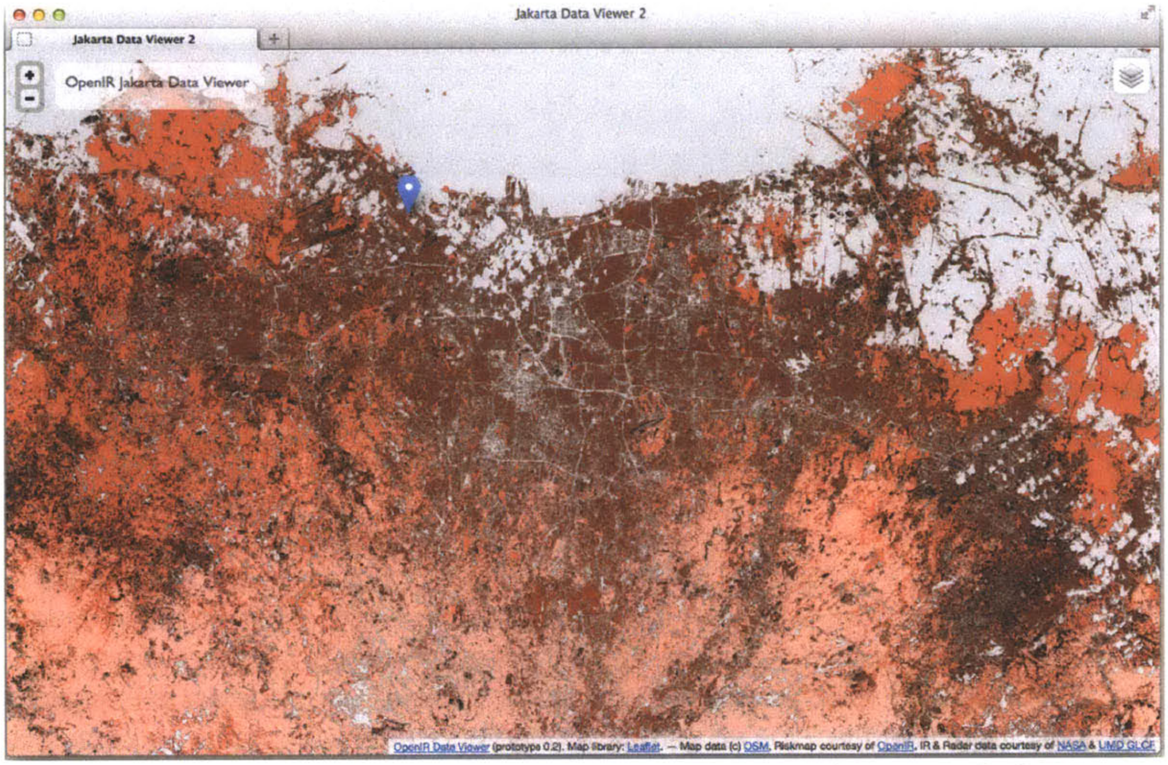
Inputs for derivation of OpenIR Risk Map. The boxed inputs have been implemented as of this writing.

After diagramming and writing out the OpenIR Risk Map algorithm, we could then write a script to call ImageMagick commands. By August 2012, the OpenIR Jakarta Data Viewer had been updated with an auto-generated Risk Index map. This map is a 4-color image that is derived from Urban, Vegetation, and Elevation inputs. This is an oversimplified approach, but I think it is enough with which to conduct usability studies and data validation:

The auto-generated Risk Index map's color key is as follows, ostensibly from highest to lowest risk:

- Deep Red: low-lying urban
- Red: high-lying urban
- Light Red: low-lying vegetation
- Pink: high-lying vegetation

With respect to environmental feature isolation, the current approach is to compile Landsat band composites, and then mask out all colors except those that indicate soil/water/vegetation/urban. A possibly more accurate approach, which I hope to take in the future, is to use spectral signatures along all available bands for any given dataset, then refer to spectral libraries (i.e. the Aster Spectral Library[51]) to identify environmental features.



Example Risk Map: Jakarta (http://openir.media.mit.edu/main/prototypes/jakarta-AutoGen).

Several technical issues persist. One concerns disambiguation. For any given image, there can be color ambiguity between data, no-data, and obstructions (i.e. cloud shadows). When all of these things are the same color, how do we disambiguate? Do we leave it to the human eye, especially for issues like obstructions?

Another issue concerns (de)contextualization. When evaluating a large image, it often makes sense to divide that area into sub-areas, evaluate those sub-areas, and then reassemble the map. But in evaluating the sub-areas, we may lose some important context. For instance, a small map chunk may not have a body of water in it so may not be evaluated as a flood-risk area, but that chunk may be adjacent to a map chunk that's part of the ocean. It seems that there's a tradeoff between spatial resolution and context.

As UNOSAT's Einar Bjorgo pointed out during his Front-End interview, no instrument is better for evaluating satellite maps and derived products than the human eye. Looking at

---

[51] ASTER Spectral Library: http://speclib.jpl.nasa.gov/

this Risk Index map prototype, there are some obvious limitations, such as ambiguity between water and cloud shadow, and limited environmental inputs.

Chapter 7: Front-End Evaluation Survey
August 2012 to October 2012

To inform OpenIR's technical development, I used SurveyMonkey to design an 8-question Front-End Survey to better understand perceptions of crowd and IR maps in communities that deal with web maps, crisis, and/or the environment. Undergraduate researcher Juhee Bae solicited responses by sending requests to several mailing lists, including:
- OpenIR
- CrisisMappers: people with a general interest in crisis and mapping
- Standby Task Force: volunteers who actively coordinate to deploy crisis maps
- Public Laboratory of Open Technology and Science: citizen scientists
- Ushahidi: a commonly used crowd map platform
- LiberationTech: people with a general interest in technology for social justice
- Humanitarian OpenStreetMap Team (HOT): people interested in or working for the HOT NGO
- MobileActive: people with a general interest in mobile
- Relevant MIT Departments: Media Lab, Department of Urban Studies and Planning, Civil and Environmental Engineering, etc.

Thirty-four people responded to the survey. Most were familiar with crowd maps and/or IR maps. Of these respondents, 94% indicated that the combination of crowd and IR maps will be useful. This response helped to validate the OpenIR project, but the detailed input of respondents indicate that this usefulness will only become a reality with significant experimentation, iterative design, and information scaffolding.

Questions on Crowd Maps

All survey respondents were asked to look at a sample crowd map picture and indicate their familiarity with crowd maps. Of the respondents, 83% indicated this kind of familiarity.

Respondents familiar with crowd maps were then asked to comment on how they are useful. Most of these comments showed the perception of crowd maps as a serious tool. Some respondents commented on particular crowd map applications, such as forest, public health, and crisis monitoring. Several respondents noted that crowd maps are powerful for aggregating and visualizing input from multiple perspectives and people, often helping users to understand various trends. One respondent even said that crowd maps "harness the power of the individual as the data maker!... often data is coming to a map from a top-down organization or government."

A crowd map's usefulness is dependent on how its categories are formed, a few comments cautioned: it is hard to convey severity of and relationships between individual reports, and categories can be key in making these connections. One comment highlighted an oft-discussed topic in the crisis mapping community[52]: the difficulty of verifying truthful information and preventing user manipulation.

There were several comments on crowd maps and public perception, with one respondent noting how crowd maps have "popularized mapping for real-time analysis in crisis response." On the other hand, a couple of respondents were quite skeptical of crowd maps, saying that the "excitement/usefulness ratio is way too high at the moment" and that crowd maps are "mostly useful for techy [sic] people who want to be important. They help influence media stories, which isn't to be taken for granted, but they provide close to zero value in on-the-ground action."

Questions on Infrared Maps

In the next section of the survey, respondents were asked to look at a sample IR image, indicate their familiarity with IR maps, and comment on the general usefulness of these maps. Of the 34 respondents, 65% were familiar with IR maps.

By far, comments on IR map usefulness focused exclusively on the natural and built environment. Again, some respondents commented on particular applications, like the monitoring of forests and other natural resources, vegetation health, illegal burning, land and water use, and weather. Several respondents noted the particular advantages of infrared imagery, noting that it is "unobtrusive," "show features more clearly than the visible light," expose the "natural landscape without the clutter of human objects," "visualize information that is not otherwise apparent," and "reveal findings that traditional aerials will hide."

Some comments addressed IR temporal and spatial resolution. One respondent wrote that IR maps allow for "rapid, unobtrusive, large area environmental classification," and another noted that IR maps help show environmental change over time. One respondent wrote that his/her familiarity with IR maps was "because of OpenIR."

---

52 Roche, Zimmerman, Mericskay: "GeoWeb and crisis management: issues and perspectives of volunteered geographic information." GeoJournal February 2013, Volume 78, Issue 1, pp 21-40

A couple of respondents offered comments that were perhaps more general or vague, including "while I don't use it, it is good to have these data available" and "every extra layer of data is potentially useful."

Finally, a few respondents mentioned IR maps in relation to the human user. One comment discussed IR maps in a larger cycle of development, saying that IR maps "help people get a macro understanding of the environment they live in, and therefore **[people] work better with their community at understanding their local impact, and [this leads to the people] doing the things to the environment which are best suited to it.**"

Questions on Combining Crowd Maps and IR Maps

When asked about the usefulness of combining crowd maps and IR maps, 94% of the 34 respondents replied positively. For this section of the survey, there were many extensive comments to questions about both the advantages and disadvantages of crowd and IR "Combined Maps."

<u>Advantages: Applications</u>

Again, many responses suggested applications for Combined Maps. Some of these suggestions were similar to those from previous sections (monitoring environmental change, disasters, etc). An increased number of responses highlighted how Combined Maps can lead to environmental justice applications, for instance:

- comparing pollutions reports with vegetation growth rates
- comparing reports of illegal activity with vegetation change and political boundaries
- mobilizing citizen scientists to use social and satellite information for corrective political action
- using IR information to help with ground-monitoring decisions
- combining IR maps with annotated findings from GIS experts

One particularly striking suggestion was public health related; the respondent suggested that a Combined Map may help with understanding why an outbreak might be happening (.e.g., lots of diarrhea near a part of a marshy lowland).

<u>Advantages: Behaviors</u>

Several respondents suggested that Combined Maps can improve social behavior in a general manner, for instance, by providing "open access info to improve oversight," showing "multi-layered effects of communities on an environment, [therefore creating] more impetus in local communities to change behaviors." One respondent suggested that

Combined Maps would change the behavior of mappers by helping to "chip away at the hegemony of pins-on-maps, just a little bit."

One respondent described a detailed scenario for economically undeveloped areas, in which Combine Maps would be "correlating future predictions of disaster with resources with need," thus changing the behavior of citizens seeking medical attention.

Advantages: Data

Many respondents commented on the advantage of one data set "cross-verifying" the other, with special emphasis on crowd-sourced data, which can be problematic in its verification. On the other hand, another response highlighted the difficulty of satellite data processing, suggesting that crowd-sourcing can help "decipher the parts [of satellite data] that are not computer readable." One respondent said that the Combined Map could be better than a street map or Google Earth for understanding "what is on the ground," while another wrote that "anytime you layer or mashup data that offers insight, it's a powerful force." This may be especially useful for faster responses to accelerating climate change, as one response pointed out.

Beyond cross-verification, respondents suggested that one dataset will inform the other, for instance, the ecological vulnerabilities revealed by IR data would be "ground-truthed" by crowd data. According to another respondent, this would give users a broader perspective; "in theory," another respondent wrote, "this should yield a really high resolution image of what is occurring."

Unsure Respondents

Several responses to both advantage and disadvantage indicated a level of uncertainty with the combination of crowd and IR data. Several respondents were simply "not sure," on the advantages or disadvantages of these kinds of maps. One wrote that he/she was not sure about using this kind of map: "I think more specific applications [of IR maps] than a crowd map would be more useful." One respondent seemed dubious as to the use of crowd maps in general: "It would be useful only to the extent that crowd maps are useful. If I have a piece of land that I need to analyze, I don't need a crowd map, I need a hyper-local map (e.g. balloon photography). If I want quantitative analysis of a huge area, crowd maps don't help."

Disadvantages: Literacy/Education

In commenting on the disadvantages of Combined Maps, most respondents noted the lack of general literacy with IR maps, and the need for greater education on how to understand IR maps. One respondent noted: "Anyone unfamiliar with IR data could easily misinterpret the data, for example by making assumptions about the colors used to represent the data. I've worked with several people who confuse IR and thermal data as well." Another respondent similarly wrote: "In my experience the resolution of satellite data is very important for people to spend time looking at it. I wonder if the colors being "different" will help or hinder people's interest in analysis."

One respondent noted that the literacy levels with crowd and IR data could complicate map usage: "People who are adding crowdsourced data might not understand the IR component to the map, and thus their contributions might not be as helpful as an expert's."

A few respondents discussed how these issues can be addressed, such as the importance of "creating an interface that novice users can understand." Another comment was more specific, emphasizing the "need [for] simple key facts and compelling data/ visualizations for people to empathize and therefore act." The multidimensionality of the data would be difficult to visualize, one respondent wrote: it would be hard to show this "data in a way that is still understandable, [and] easily browsable."

<u>Disadvantages: Data</u>

Several respondents saw limitations to the way that crowd and IR data can verify each other. Considering the rigorous way that satellite images are verified and corrected, some respondents noted that a similar level of oversight would be needed for crowd map data to be scientifically useful. High sample sizes of crowd responses, noise reduction, and training would all be needed for crowd data to be more useable.

Respondents noted other difficulties in combining crowd and IR data, including bandwidth demands, and the "difficulty of selecting appropriate, common parameters" between data sets. One respondent expanded on the notion of differing scales between crowd and IR data, saying that the combination is "not helpful for broad research, not helpful for hyper-local analysis, and not as easily accessible to the media as pins-on-maps."

Takeaways

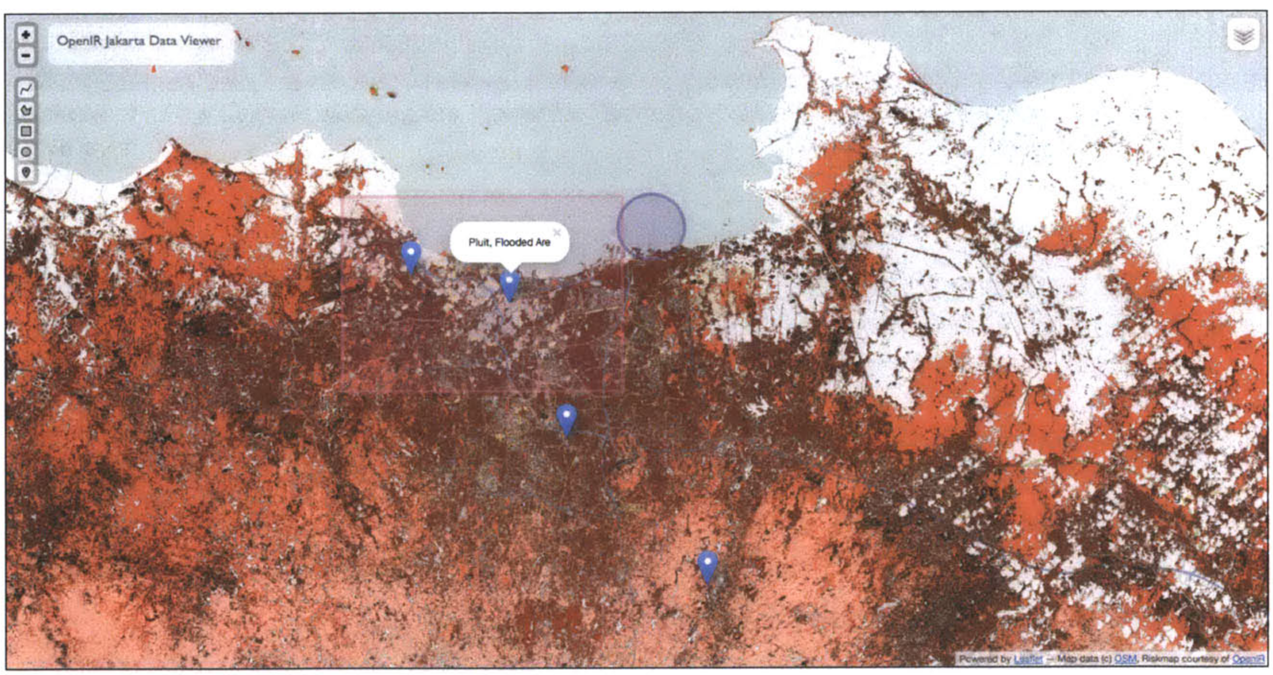
This evaluation showed that there is significant interest in the combination of IR and crowd data. It may lead to practical environmental applications, help to encourage more environmentally conscious behavior, and lead to multifaced insights through data "cross-verification."

However, there are major concerns as to this combination being understandable or usable. This stems in part from large disparities of spatial scale and truth validation between IR and crowd datasets; it also stems from a lack of general public literacy around IR imagery. In developing our data viewers for testing, we attempted to address some of these concerns with scale, validation, and literacy.

Chapter 8: Implementation Phase 3
September 2012 to November 2012

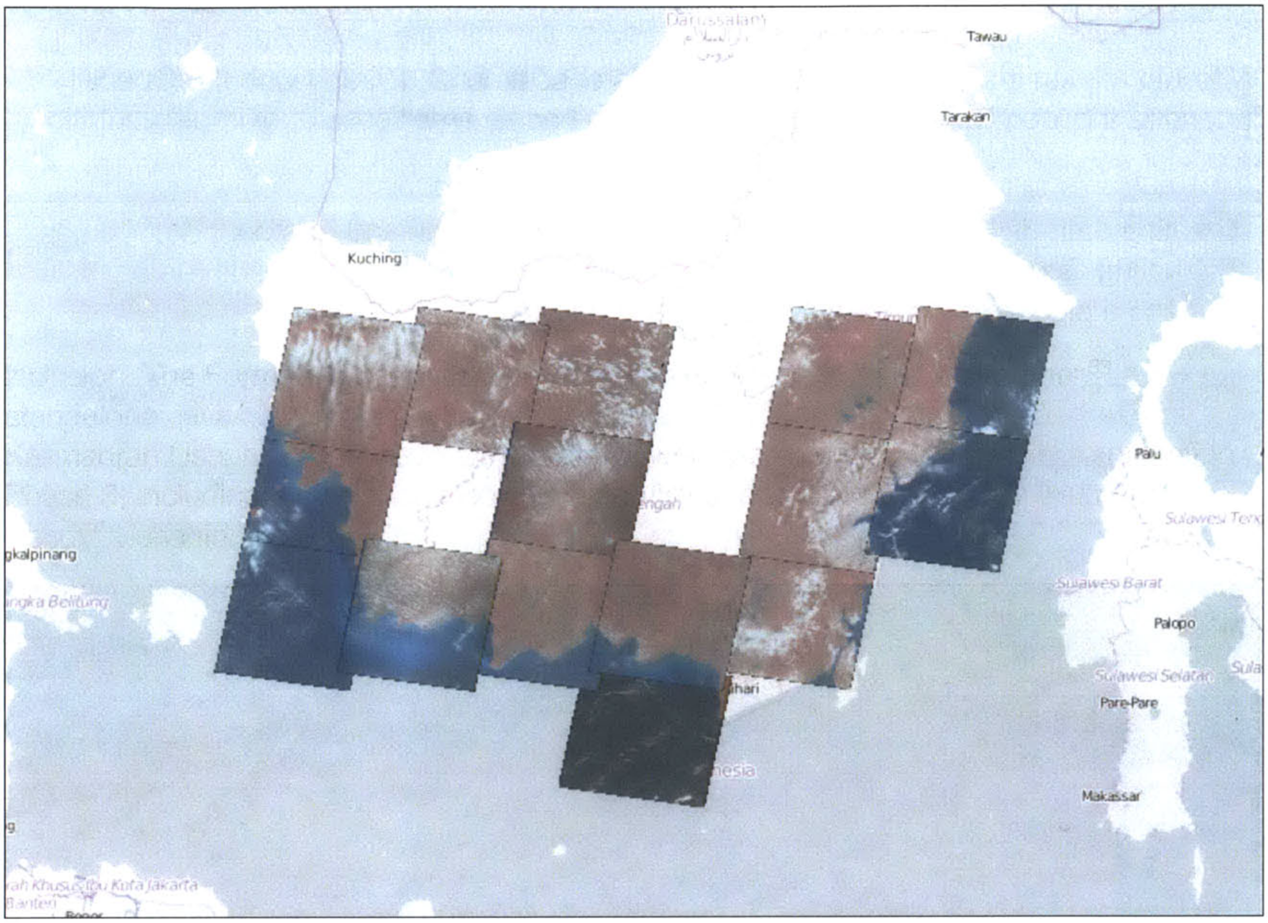
With information from the Front-End interviews of summer 2012, and with the OpenIR Indonesia trip looming ahead for January 2013, I had an ambitious set of implementation goals for autumn 2012. They included

- Creating data viewers for new study areas: Java/Kalimantan and Boston.
- Evaluating Google Earth Engine.
- Building a plug-in to include OpenIR layers in an Ushahidi deployment.
- Developing a set of drawing tools for the data viewer user interface.
- Implementing analytics code to track user interactions with the OpenIR Data Viewer maps.
- Experimenting with color and spectrograph functionalities for the data viewers.



Screenshot of the OpenIR Jakarta Data Viewer for Formative Evaluation, with the Risk Index Map as the active layer.

The features resulting from these goals were then integrated into OpenIR Data Viewers used for Formative Evaluation interviews with seven stakeholders (see Chapter 9) The only exception were color and spectrograph functionalities, which were too early-stage for inclusion. The Formative Data Viewers were a standalone Jakarta Data Viewer,[53] a standalone Java/Kalimantan Data Viewer,[54] a Jakarta Ushahidi instance,[55] and a Java/Kalimantan Ushahidi instance.[56] Other data viewers were built during this Implementation Phase 3, including a standalone Boston Data Viewer and a New York City "Hurricane Sandy" Ushahidi instance.



Screenshot of incomplete OpenIR Java/Kalimantan Data Viewer, with the Vegetation layer on display. The completed Formative Data Viewer for Java/Kalimantan includes band combinations for all of Kalimantan.

New Study Area: Java/Kalimantan

Kalimantan (Borneo) is one of Indonesia's largest islands, even though it is shared with Malaysia and Brunei. Still largely covered in tropical forest, it is the focus of much local and

[53] http://openir.media.mit.edu/main/prototypes/jakarta_drawing2/example/index.html

[54] http://openir.media.mit.edu/main/prototypes/jakarta_multiple_tiles/example/index.html

[55] http://openir.media.mit.edu/hackathon/Ushahidi/

[56] http://openir.media.mit.edu/Ushahidi/Java_Kalimantan/index.php/

international environmental activity, and the center of many land use issues between indigenous tribes and corporate plantations. OpenIR community partner Harry Surjadi spent two International Knight Journalism Fellowship periods in Kalimantan, developing citizen journalism-styled systems with indigenous advocacy groups in the provincial capitals of Pontinak (West Kalimantan province) and Palankaraya (Central Kalimantan province). Surjadi calls these "Information Broker" systems, in which local citizens share information like citizen journalists, and they use the information to "broker" change with governments, companies, and other groups.

OpenIR was interested in working, under Harry's guidance, with these groups and perhaps connecting their Information Broker systems to OpenIR's own. To do this, data viewers for Kalimantan had to be built. Since OpenIR also planned to meet with groups in the city of Yogyakarta, which like Jakarta, is on the island of Java, my collaborator Ilias Koen and I developed a data viewer to view processed IR satellite images (known as "tiles") for all of Java and Kalimantan. Ilias ran scripts to convert the imagery from GeoTIFFS to Tile Map System (TMS) image sets, while I built a custom data viewer.

While the conversion was not complicated, it was still a processor- and time-intensive task that included converting dozens of tiles at fourteen "zoom levels." A user changes zoom level whenever he/she clicks the "+" or "-" button of a web map, usually in the upper-left area. With TMS maps, each zoom level translates to a directory containing a PNG subsection of the entire map. The higher the zoom level, the more PNG subsections; for instance, zoom level 5 may have a few PNG subsections for one Landsat tile, while zoom level 14 may have thousands of PNG subsections for the same tile.

For the Java/Kalimantan viewer, the data conversion process was split into three major scripts:

- getTilewithPathAndRowOnly downloads all the Landsat bands for a single tile. The bands are in zipped GeoTIFF format.

- createTMSforSingleBandWithDirectories calls the Geospatial Data Abstraction Library to generate TMS web files for single (greyscale) bands of tile mosaics. The generated files include directory structures, HTML and JS files, and PNGs of the sliced map.

- createTMSCombinationWithDirectories calls ImageMagick command-line software[57] to make band composites of the sliced PNGS. These band composites are then used to make a new data viewer that contains OpenIR's standard band composites: 321 (true color), 432 (vegetation), 543 (soil), 453 (urban), and 754 (moisture). By default, the script only processes zoom levels 5-7. But by adjusting the maxZoom and minZoom variables, the script will process all zoom levels from 1 to 14. However, the time to process higher zoom levels is much, much longer.

---

[57] ImageMagick: Convert, Edit and Compose Images. http://www.imagemagick.org/script/index.php

This process is an update on the ImageProcAndPrep process developed during Summer 2012 (see Chapter 6), in that the GeoTIFFS are first sliced/diced into PNGs, and then the resulting PNGs are composited into the colored images. There are a few advantages to switching these steps:

- It is faster and more modular to composite lots of little images instead of one huge image.
- The original greyscale PNGs are preserved, so they could potentially be combined for other web maps. Our thinking is that we may collect these PNGs on one server so that they are available for other kinds of band combinations.

With its fourteen zoom levels, the Java/Kalimantan data viewer became OpenIR's most extensive and heaviest (~12 GB) data viewer. The process of developing it highlighted one unsolved problem in our approach to IR datasets: the lack of color and atmospheric correction. Without this kind of correction, conducted automatically or manually, the IR tiles have a dull, muted color. In OpenIR's single-tile data viewers, it's possible to manually color-correct the one tile, but for a huge data viewer like Java/Kalimantan, manual processing is not an option. For the time being, all the image colors in this data viewer remain uncorrected.

New Study Area: Boston

In spring 2012, I entered a competition to win ten MudWatt microbial fuel cells (colloquially called "mud batteries") from the company KeegoTech. I won this competition by writing a lesson plan about different kinds of muds and their conductive abilities to power mud batteries. Prior research has found that agricultural soil is better than forest soil at powering mud batteries.[58]

I wanted to experiment with a group of batteries, each containing a different soil sample from one metropolitan area, as an in-situ network from which soil data would be visualized in an OpenIR data viewer. Similarly to how I hypothesize that IR satellite and crowd data could inform each other, I was interested in how ground-based soil data could inform findings from satellite-based soil data. Since I currently live in Cambridge and have the most access to its soil, I built an OpenIR Data Viewer for the Boston area. I collected several samples, including some from farms, swamps, cities, and forests, and plotted their sample locations on the Boston data viewer. I filled each battery with a soil sample and waited for the battery to start powering up, as indicated by a simple circuit and LED wired

---

[58] Sara J. Dunaj, Joseph J. Vallino, Mark E. Hines, Marcus Gay, Christine Kobyljanec, and Juliette N. Rooney-Varga. "Relationships between Soil Organic Matter, Nutrients, Bacterial Community Structure, And the Performance of Microbial Fuel Cells." Environ. Sci. Technol., 2012, 46 (3), pp 1914–1922

to the battery. My aim was to build a system, perhaps based on the Twine environmental sensing device [59] to collect power, current, and resistance readings of the mud batteries, and then demonstrate this network, using the OpenIR visualization, at the MIT Media Lab.

Unfortunately, no battery ran for more than one month, and after some simple troubleshooting, I dropped this admittedly peripheral project. It was still useful to build a Boston data viewer and demonstrate some of the OpenIR+MudWatt concepts at the World Maker Faire in New York. This data viewer was also useful for other Media Lab colleagues who wanted a local version of OpenIR.

Screenshot of OpenIR Boston Data Viewer, with the Soil layer on display.

Google Earth Engine

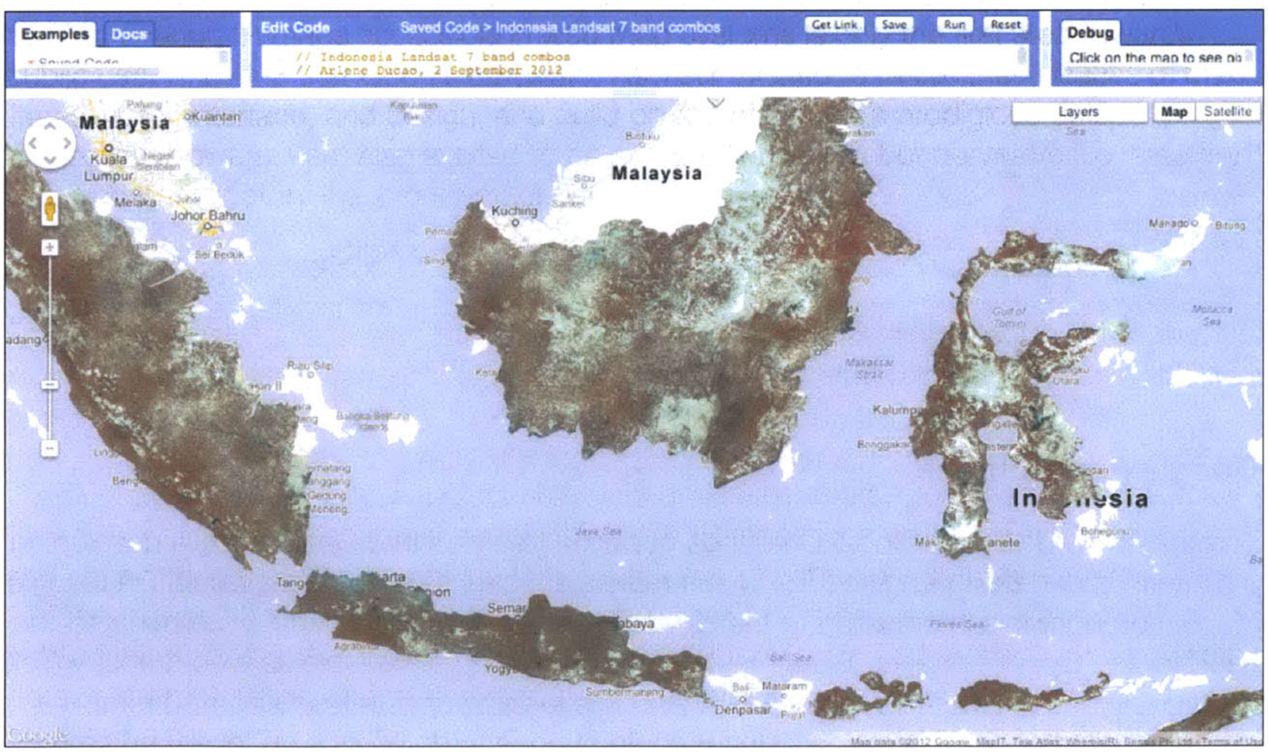
Through the MIT IDEAS Global Challenge network, I learned that Google.org is developing a platform called Earth Engine (EE), which makes spectral satellite data available in Google Maps. I brought this information to the OpenIR team, which wondered if EE eliminated the need for the OpenIR project. EE is making most free IR data available as web maps, which is precisely one of OpenIR's major goals. And it is tackling spectral, spatial, and temporal resolution all at once, something OpenIR couldn't hope to do on its own for a long time.

---

[59] Smalley, Eric. "'Twine' Seeks To Tie Up The Smart Environment." Wired Magazine: November 29, 2011. http://www.wired.com/business/2011/11/twine-cloud-smart-environments/all/ Visited April 2013.

On the other hand, the availability of EE could catapult OpenIR several steps down the path to its longer-term goals of increased literacy, creative usage, and better user interfaces with IR satellite data. A closer look at the EE web site reveals that there is still much work left in making this data usable for general audiences. For one thing, the site quite rightly specifies EE as a tool for scientists and researchers. Considering that IR map literacy is not widespread (most people think that IR maps look weird or Photoshopped), and considering that IR maps of the Earth didn't even exist before satellite systems of the 1970's, making IR maps accessible in web map systems is only the first step towards public use. Knowing that Google EE exists, I am still excited to tackle these issues of map literacy, user interface, and design, and build on top of Google's prodigious data processing powers. I am also excited to be at the meeting point of pervasive (i.e. satellite) data and participatory (i.e. crowd-map) data.

In August 2012, my collaborators Ilias Koen, Aziz Alghunaim, and I were added to Google.org's Earth Engine's Trusted Tester program, and we started testing EE's APIs in early September. EE offers some of the freely available datasets most widely used in remote sensing work: Landsat satellite bands, MODIS satellite bands (daily, 8-day composites, and 32-day composites), and MODIS derived products like NDVI (vegetation classification). Many of these products are explained and visualized on the EE site.



Screenshot of Google Earth Engine, Appspot "sandbox" development environment. On display is a Landsat 432 band combination for Indonesia. This is the same band combination used in OpenIR's "Vegetation" layers.

EE is currently available as a secure Javascript and secure Python API. The Javascript library is only usable in an Appspot.com sandbox, and the Python library can be downloaded locally and used, once an OAuth2 authentication is set in place. We experimented with calling the Python API from our server via a Flask-based proxy server, but the API is not complete enough to use with web map libraries. As a result, the only currently feasible methods for using EE are in the Appspot.com sandbox or on a local machine. When the Javascript library's authentication procedures are documented, it may be feasible to call the EE Javascript API remotely. When this becomes the case, we will start to include EE in OpenIR data viewers. In the meantime, it has been useful to process the data independently, so as to be less dependent on a project that may be shut down at any time.

In a November 2012 meeting with Dave Thau, Goggle EE Developer Advocate and OpenIR's technical adviser, my collaborators and I discussed whether EE obviates the need for OpenIR. Thau pointed out that EE does not focus on general user experience or accessibility; rather, EE is marketed to geospatial professionals. While EE's API may become more open in the coming years and months, this target market is not likely to change. OpenIR's studies with general users on notions of literacy, experience, and interaction with remote sensing, are thus useful ways to expand the data availability afforded by EE's APIs.

Ushahidi Plug-in

In the past five years, Ushahidi has become one of the most widely used crowd-map platforms, that is, a platform that allows any user to contribute his/her geo-located communications (i.e. e-mail, SMS, Tweet, web-form submission), which are then automatically pinpointed on a web map. While TMS raster imagery has occasionally been incorporated into Ushahidi deployments, its incorporation tended to be unsystematic and ad-hoc.

While the Ushahidi system is not the only open-source crowd-map in existence, and it may not be the most elegantly designed, it is the product of a thriving open-source community—not only do a lot of people use it, a lot of developers contribute to its source code. It is also widely used for tracking environmental crises, including the 2010 Haiti earthquake and the 2010 BP oil spill off the coast of Louisiana. For all of these reasons, I thought it made sense to systematize the incorporation of TMS imagery into Ushahidi by developing a plug-in for the system. Heather Leson, Ushahidi's Director of Community Engagement, was very enthusiastic about this idea, and she connected our team to the organization Geeks without Bounds (GWOB), which holds coding marathons to build humanitarian tools. Our team participated in two GWOB "hackathons," and this was enormously helpful in accelerating the development of an Ushahidi/OpenIR TMS plug-in.

GWOB Hackathon 1: MIT, October 5-6, 2012

This event was sponsored by AT&T and MIT Sloan Business Club and was held at MIT's Stata Center. GWOB's director, Willow Brugh, kindly offered an "OpenIR Prize" at this event: $1000 to be split amongst contributors, not already on the OpenIR team, who worked on the Ushahidi/OpenIR plug-in.

To prepare for the event, undergraduate researchers Stephanie New, Aziz Alghunaim, and I set up a "hacker sandbox" on the OpenIR server—an isolated FTP partition in which outside contributors could add their code, without the worry of this code affecting other parts of the OpenIR system. In this sandbox, we deployed an Ushahidi instance, which includes

- setting up a dedicated MySQL database
- downloading the Ushahidi code [60]
- modifying permissions on certain files so that user records can be written to them
- running the Ushahidi PHP installer.

To the sandbox, we also uploaded a copy of the TMS layers from the OpenIR Jakarta Data Viewer. I then wrote a "hacker brief,"[61] to explain to contributors how they should use the sandbox and external resources to try to develop a Ushahidi/OpenIR plug-in. Part of why I wrote this brief was to better organize the limited web documentation on incorporating TMS imagery into Ushahidi.

On the day of the GWOB MIT event, four contributors joined the OpenIR team. The first thing that we examined was how the Ushahidi platform called maps from the map servers from the providers ESRI, Google, Bing, and OpenStreetMap. The team decided to focus on the ESRI imagery layer because it was calling its map in tiles (subdivided imagery), and OpenIR's Jakarta map is also partitioned into tiles. ESRI was also very promising because there was no API call (OpenIR has not yet created an API to communicate with other software).

We created a new layer called OpenIR in Ushahidi's "map.php" file, emulating the structure of the ESRI Imagery layer.

```
// OpenIR Imagery
$layer = new stdClass();
$layer->active = TRUE;
$layer->name = 'openir_imagery';
```

[60] Download the Ushahidi Platform. http://download.ushahidi.com/

[61] OpenIR Hacker Brief 1: GWOB. http://openir.media.mit.edu/main/?p=667

```
$layer->openlayers = "TMS";
$layer->title = 'OpenIR Map';
$layer->description = 'This is the OpenIR layer';
$layer->api_url = '';
$layer->data = array(
'baselayer' => TRUE,
'attribution' => '',
'url' => 'http://openir.media.mit.edu/main/prototypes/jakarta/Tiles_Jakarta_Riskmap/${z}/${y}/${x}',
'type' => ''
);
$layers[$layer->name] = $layer;
```

For the OpenIR layer, we changed the tile type from 'XYZ' to 'TMS,' since OpenIR's Jakarta layers are in the TMS (Tile Mapping Service) format, not the XYZ (basic grid) format.

```
$layer->openlayers = "TMS";
```

We then focused the most on the server URL call. Two kinds of URLs were tried:

```
'url' => 'http://openir.media.mit.edu/main/prototypes/jakarta/Tiles_Jakarta321/${z}/${y}/${x}',

'url' => 'http://openir.media.mit.edu/main/prototypes/jakarta/Tiles_Jakarta321/${z}/${y}/${x}.png'
```

All of these URL failed to call the tiles on the OpenIR server; on the front end, the map console in Ushahidi displayed broken image icons instead of OpenIR layers.

We tried a hard-coded OpenIR server call:

http://openir.media.mit.edu/main/prototypes/jakarta/Tiles_Jakarta321/10/817/494.

With this URL, a tile was properly displayed in the Ushahidi map console. The server url seemed to be working, but the OpenIR layer would not show up in Ushahidi without a hard-coded file location.

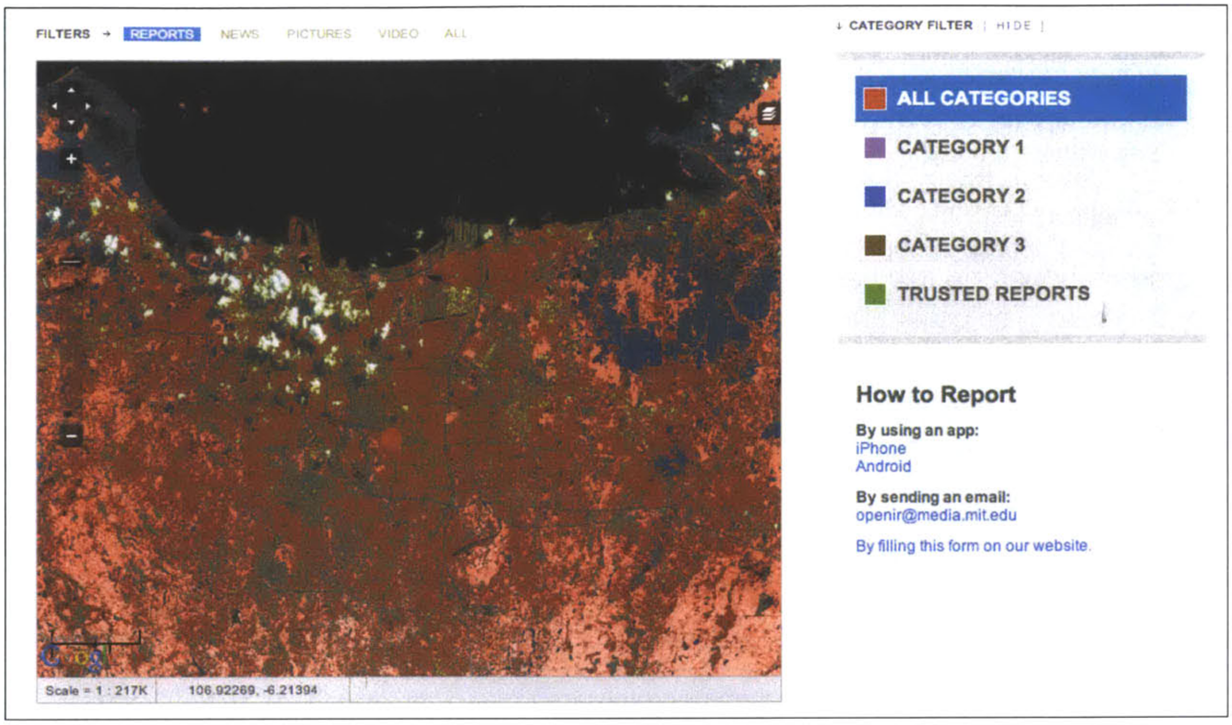
Although our team did not emerge from the GWOB event with a working plug-in, it was still a great motivator, and I've come to believe that "hackathons" can be a useful way to jump-start a small project. After this event, Ilias tried converting an existing Ushahidi plug-in, Seth Kigen's WMS (World Map System, another web map protocol) plug-in[62] and started to have better results with TMS imagery in Ushahidi.

GWOB Hackathon 2: George Washington University, October 13, 2012

This event was held as part of the International Conference on Crisis Mapping (ICCM) in Washington D.C. Both Heather Leson and Willow Brugh were present, as was software engineer Chris Willard, who joined the OpenIR team during this event.

---

[62] Ushahidi-plugin-wms. https://github.com/kigen/Ushahidi-plugin-wms



Screenshot of OpenIR/Ushahidi build after the GWOB ICCM event.

During this event, path naming was again a primary issue. Ushahidi's bundled OpenLayers, a javascript mapping library, was inserting two directory names into the image paths we specify. For instance, if we specified the name "http://hoppy.com/7/5/165.png," OpenLayers would send "http://hoppy.com/serviceVersion/layername/7/5/165.png" to Ushahidi.

This "serviceVersion/layername" could not be removed without re-compiling OpenLayers, but they can be controlled with this new bit of code in the OpenIR/Ushahidi plug-in directory "helpers":[63]

```
return $var_name ."= new OpenLayers.Layer.TMS( \"$layer_name\" , \"http://dukodestudio.com/openIR_tileMaps_Wdc/$layer_title/\", { serviceVersion:'1.0.0', layername: '754', type: 'png'} );
```

In the line above, the values for 'serviceVersion' and 'layername' can be changed or removed. We referred to the the OpenLayers TMS script[64] to help solve this problem. After this, we could successfully include our OpenIR TMS layers in an Ushahidi deployment. As of now the OpenIR/Ushahidi plugin, for internal use only, is on Github.com.[65] There is no UI

[63] Example Helpers directory: https://github.com/kigen/Ushahidi-plugin-wms/tree/master/helpers

[64] OpenLayers: TMS.js: https://github.com/openlayers/openlayers/blob/master/lib/OpenLayers/Layer/TMS.js

[65] DuKode: OpenIR_ushahidi: https://github.com/DuKode/openIR_ushahidi

for uploading and replacing TMS layers; this is modified in the plugin code—this is why I think the plug-in is only usable internally at this point.

Hurricane Hackers, MIT, November 3, 2012

The GWOB events motivated my team and me to build a functional plug-in for Ushahidi. However, our Ushahidi deployment was not set up to be very usable or useful. Hurricane Sandy, an anomalous autumn "superstorm" that seriously damaged New York City and the Jersey Shore, provided our team with an opportunity to build more usable, useful combinations of OpenIR and Ushahidi. This event was my first encounter with the tensions of building technology for environmental disasters: on the one hand, the actual disaster can be a highly useful learning opportunity, but on the other hand, the disaster causes a lot of human suffering. In the case of Hurricane Sandy, some of that suffering was my own, as I own a small Brooklyn business that is located in a hurricane evacuation zone.

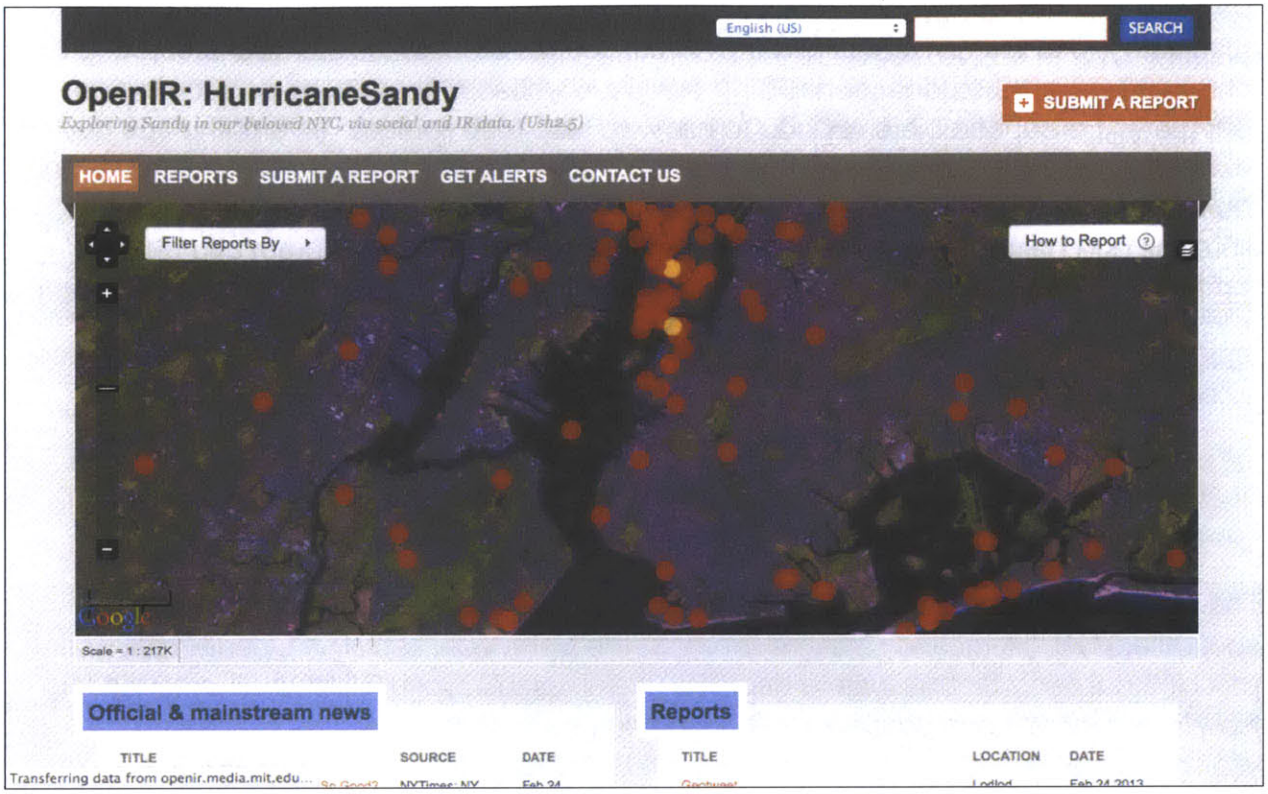
The weekend after Hurricane Sandy, some colleagues in MIT's Center for Civic Media spearheaded a "Hurricane Hackers" event at the MIT Media Lab. The OpenIR team participated, and our goal was to deploy a useful, usable OpenIR/Ushahidi instance for New York City. We had processed IR band combinations for New York City earlier that year, so I thought it was feasible to set up an Ushahidi system that could still be usable and visually useful, even though the storm had passed.

First, we set up an Ushahidi build with the newly released 2.6 platform. However, our internal OpenIR plugin, which worked so well with Ushahidi 2.5, didn't work in v.2.6. We're starting to work on fixing this, but in the meantime, we installed and set up our TMS layers in v.2.5. We also updated settings in the Ushahidi dashboard, including adding RSS newsfeeds, changing the appearance of the deployment, etc.

Since the hurricane was over, we thought that few people would want to actively submit reports to our deployment. So to populate the map with social data, we integrated geotagged tweets with the hashtags #Sandy, #Hurricanesandy, and #Frankenstorm, into the map.

To automatically turn a geotagged tweet into a report in the data viewer, we first had to choose which tweets were relevant. We picked the hashtags #Sandy, #Hurricanesandy, and #Frankenstorm in the Settings tab of the admin's Dashboard. We then created a trigger in the "experimental" Actions section of the Ushahidi dashboard.

This Ushahidi build was more of a research simulation than an active tracker; it was made too late to be used during Sandy, although it is currently collecting some useful geotweets about recovery needs. As the tweets start to populate the map more densely, we will study how they cross-reference land features highlighted by the IR data.



Screenshot of OpenIR/Ushahidi build for Hurricane Sandy.

This Hurricane Hacker event, combined with the previous GWOB events, very much helped my team and me learn to make a useful Ushahidi deployment with OpenIR layers. After these events, the OpenIR team could compile Landsat IR combinations, build them into an Ushahidi deployment, and add geo-tweet feeds—things that we can now do at the beginning of future disasters. These skills turned out to be very useful during our time in Jakarta, which was severely flooded on January 16, 2013.

Data Viewer Drawing Tools

The OpenIR Data Viewers are not yet fully customizable: users cannot yet individually modify, save, and export maps within the OpenIR system. To provide a rudimentary level of customization, undergraduate researcher Stephanie New and I added a set of simple drawing tools to all of our data viewers starting in October 2012. With these tools, the user can now draw lines, polygons, rectangles, and circles. A pin/marker with a user-defined popup message also can be added to the data viewer.

These drawing tools are a modified version of Leaflet.draw, Jacob Toye's vector drawing plug-in for the Leaflet map library.[66] OpenIR's primary contribution is the customization of pin/marker messages. This modified version is available through GitHub.[67]

Back-End Analytics: Map Events

User feedback through surveys and interviews provides greater understanding of how OpenIR is and can be used. Supplementing this information with passive, automated data collection can provide a much more rounded picture of OpenIR's use. To this end, undergraduate researcher Srinidhi Viswanathan and I implemented code to track a user's interaction with OpenIR Data Viewer maps. For instance, changing a map layer, dragging it, zooming into it, drawing shapes on it—all these events would be tracked for each user and recorded, with the user's IP address used to identify and differentiate between users.

Since OpenIR's Data Viewers are currently based on the Leaflet map library, Leaflet tracking commands are used to trigger a series of actions. In an early implementation, a Leaflet event would trigger a Leaflet alert box to appear and a record to be written, via an ActiveXObject, to a text file.[68] However, this implementation led to several issues: the alert box was irrelevant for the user, the ActiveXObject would only work in Internet Explorer (as opposed to Google Chrome, Apple Safari, Mozilla Firefox, and other web browsers), and the text file could not be manipulated like a table.

To address these issues, we implemented these changes:

- The text of the alert box was output to the server rather than the client-side. The alert box was disabled on the browser; it was simply used for data-gathering purposes
- Javascript was used to track the various events being clicked on the website. Then PHP was used to track this process on the server. We also implemented AJAX (asynchronous javascript and XML), which ensures that response is not delayed when a layer is clicked, thus enhancing the web interaction.
- Our initial code tracked only additions and removals of layers to a map. Later, we expanded this tracking include layer dragging, zooming, and drawing. After this expansion, each icon clicked in an OpenIR data viewer is tracked on the UpdateServerLog file on the server side in a CSV (comma separated value) format.

A sample CSV snippet is below. It follows the format
EVENT CATEGORY, user IP address, date, time, event

---

[66] Leaflet.draw. https://github.com/jacobtoye/Leaflet.draw

[67] DuKode: OpenIR_drawing. https://github.com/DuKode/OpenIR_jakarta_drawing

[68] ActiveXObject (Javascript). http://msdn.microsoft.com/en-us/library/ie/7sw4ddf8(v=vs.94).aspx

```
LAYERADD,50.53.49.234,2012/12/12 23:35:53,Moisture OFF
LAYERADD,50.53.49.234,2012/12/12 23:35:56,Soil OFF
DRAW,50.53.49.234,2012/12/12 23:36:06,Rectangle
LAYERADD,50.53.49.234,2012/12/12 23:36:13,Soil ON
DRAW,50.53.49.234,2012/12/12 23:36:20,zoom level 12
DRAW,50.53.49.234,2012/12/12 23:36:20,drag
```

At the conclusion of the Summative Evaluation period in March 2013, we will analyze these CSVs in tandem with the Summative surveys and interviews.

Color/Spectrograph Experiments

A significant advantage of infrared band combinations is their use of color to clearly and vibrantly highlight environmental features in a map. For instance, the combination of Landsat bands 4 (0.76-0.90 micrometers), 3 (0.63-0.69 micrometers), and 2 (0.52-0.60 micrometers) is widely used for vegetation studies: in the resulting images, healthy vegetation is clearly colored in deep red hues, while other environmental features are downplayed in contrasting, duller hues.

OpenIR's long term innovations will likely address user interface techniques for translating color to environmental information. In November and December 2012, Ilias Koen, Stephanie New, and started to experiment with these kinds of techniques.

Our first experiment involved the "eyedropper" metaphor common in image processing software like Adobe Photoshop: when a user clicks on a screen-pixel of a desired color, the color of that pixel is recorded for future use. In applying this approach to the javascript-based OpenIR data viewers, we encountered this hurdle: an image object had to be instantiated before the color of a chosen pixel could be returned. With the draggable IR layers in the data viewer, declaring a static image object was not feasible.

Our next experiment was more ambitious, and potentially much more powerful. We built a small application to display a spectrograph for a clicked pixel in any given Landsat image. In this application:

- The user clicks on any pixel on the map.

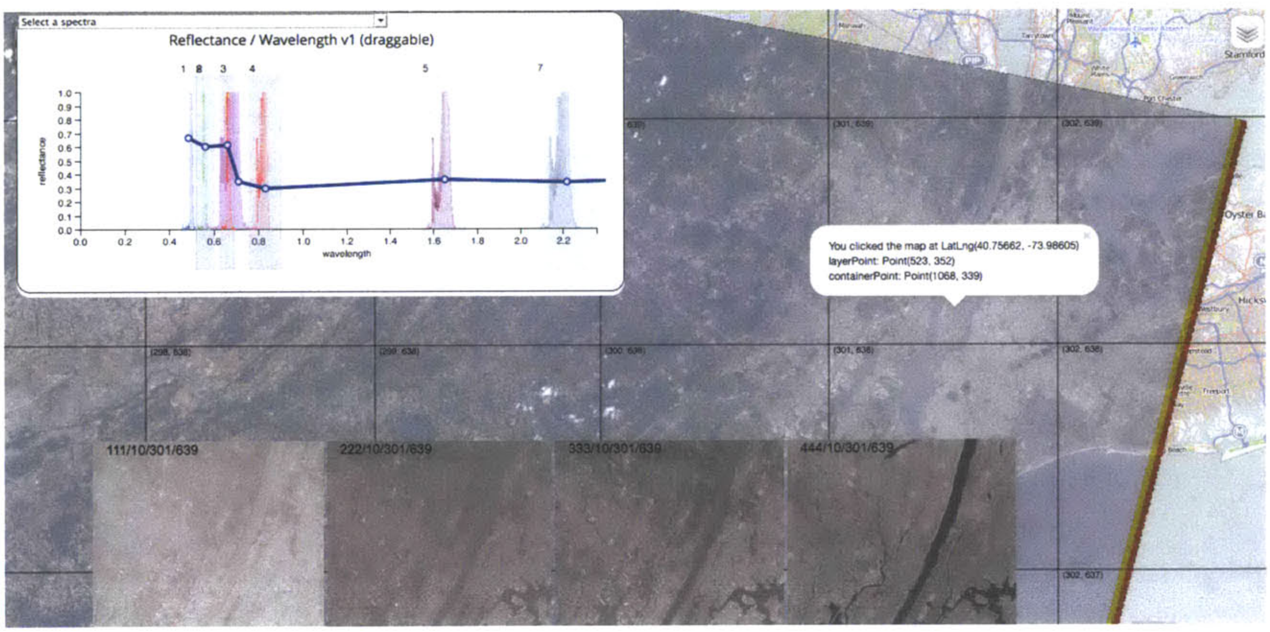
- The latitude, longitude, zoom level is used to calculate the surface reflectance (pixel luminance) for each band (wavelength range).

- This information is plotted as a histogram, which shows a combined surface reflectance signature for the pixel.

- The histogram is compared with the ASTER spectral library, which is an extensive index of environmental features (i.e. "construction pavement," "dry vegetation") for surface reflectance signatures.

- If an environmental feature is identified for the pixel's spectral signature, it will be displayed as a pop-up. In this manner, clicking on a pixel will yield a statement of environmental detail, i.e. "You clicked on construction pavement," "You clicked on dry vegetation," etc.

While the spectrograph app is currently too rough for inclusion into our data viewers, its environmental implications are wide ranging. For instance, an algorithm based on the spectrograph app could open, to the general public, environmental details like building age, construction site materials, vegetation health, and water pollution. Just as true-color satellite imagery, once under high security, is now common in web-maps, IR environmental details, still esoteric and also often a kind of high-security information, may become more ubiquitous in the coming years.



Screenshot of Spectrograph application.

Takeaways

This implementation period was the most intense since OpenIR's launch. While the Indonesia trip that followed this period would teach us a great deal, this implementation period left us with a few important lessons.

- Hackathons were of great value in accelerating progress towards a pre-defined goal. Some developers bemoan how hackathons have devolved away from their more free-spirited, brainstorming-oriented origins[69], but I nevertheless found it very useful to approach the hackathon as an intense coding session with the fresh new input of students, crisis responders, and developers.

- Hurricane Sandy gave our team a small taste of how deploying an experimental crisis-related technology in a time of crisis is both gratifying and dispiriting: gratifying because a contribution is being made, but dispiriting because the contribution is so minuscule, especially in the face of such destruction. This feeling would be even more intense during the Jakarta monsoons in January 2013.

- Data Preparation over large regions, like in the case of Java/Kalimantan, is still very effort intensive for a team with our resources. On the other hand, streamlining data preparation continues to be an extremely useful exercise that allows us greater control over image clearness and vividness. Experimenting with Google Earth Engine gave our team many ideas for how an IR data API would streamline OpenIR's efforts to make IR data more available.

- This implementation period consisted mostly of small, disparate experiments that resulted in a disparate data viewers. While overall, it was useful to develop many features in their own "sandboxes," these features have not yet been combined into one showcase application. It may have led to stronger user response if this combination had been developed before the Indonesia trip.

[69] Personal conversation with Chris Willard, January 2013.

Chapter 9: Formative Evaluations
October 2012 to January 2013

After completing training from MIT's Committee on the Use of Humans as Experimental Subjects (COUHES), I submitted a set of Formative and Summative interview questions, as well as Summative survey questions, for approval to COUHES. These questions require very little personal or private data from respondents, and they were approved without rebuttal.

The decision to evaluate OpenIR's user interface through a series of Formative and Summative evaluations stemmed from a discussion of "thesis types" during the MAS Thesis Class, Autumn 2012. Professors Pattie Maes, Chris Schmandt, and Mitch Resnick explained three general categories of MIT Media Lab theses: science/engineering, art/design, and human-centered. The evaluation process for a human-centered thesis, it was explained, consisted of rounds of Front-End, Formative, and Summative evaluations with users.

I was familiar with this evaluation process through years of working with informal education institutions, namely the American Museum of Natural History (AMNH) in New York, which is required to conduct these kinds of evaluations with exhibits and media funded by federal grants. In fact, there are a handful of companies that specialize in Formative and Summative evaluations for federally-funded media.

I also thought that this approach made much sense for OpenIR's evaluation process, because though there are strong design and engineering aspects to the project, I think its long-term success hinges on its human-centered goal of increasing environmental data accessibility and literacy to the general public. OpenIR's Data Validation processes help the team better understand its engineering design, while its Formative and Summative evaluations help the team better understand its human-centered design.

OpenIR's Formative and Summative interview and survey questions are based on similar evaluation materials used by AMNH for recent interactive exhibitions and animations.[70] The AMNH materials were created by companies including the Institute for Learning Innovation. For OpenIR, the questions were modified to adequately evaluate the workings of a web application as opposed to a museum exhibit.

---

[70] American Museum of Natural History: Evaluation. http://www.amnh.org/learn-teach/evaluation-research-and-policy/evaluation

For the Formative evaluations, 30-60 minute structured interviews were held with stakeholders in the Indonesia and international mapping communities. Some of these respondents were also OpenIR community partners. The Formative evaluation respondents were (in order of interview date):

- Harry Surjadi, environmental journalist (Indonesia)

- Kate Chapman, Acting Director, Humanitarian OpenStreetMap Team (Indonesia)

- Desiree Matel-Anderson, FEMA (U.S.)

- Suwandi Ahmad, AirPutih (Indonesia)

- Christopher "Scott" Shoup, FEMA (U.S.)

- Trias Aditya, Universitas Gadjah Mada (Indonesia)

- Stefanus Masiun, Ruai TV (Indonesia)

Other than helping to orient undergraduate researcher Juhee Bae during the first Formative interview, I did not participate in these interviews, so as not to influence the evaluation. Part of this rationale is that I am both the lead coordinator and a software programmer for the project. Juhee is OpenIR's usability and community researcher; having not directly participated in design or implementation, she was a more neutral party to conduct these Formative interviews, which she did using Skype.

Feasible Points to Address (Summative Data Viewers)

Formative interview questions were approved by COUHES in October 2012. Because of time zone differences and the busy schedules of all involved, our team continued to schedule Formative interviews until early January 2013, immediately before leaving for Indonesia to conduct fieldwork and move to the Summative stage. This gave the team only a short time to analyze the Formative interview results and construct new Summative Data Viewers for evaluation during and after the Indonesia trip. To help with this, Juhee highlighted interview points that were both relevant to OpenIR's Indonesia community partners and feasible to rapidly address in a few design+coding sessions:

- including legends/keys (this was the most suggested response)
- including a transparent layer with street names (this was the second-most suggested response)
- implementing hover-based instead of click-based explanations of environmental features
- allowing Risk Index levels to be toggled separately

- allowing users to see the layers when zooming in (some zoom levels are not processed in the Jakarta Data Viewer)
- ensuring the responsible use of social media when adding tweets to Ushahidi map
- using Indonesian words for Twitter hashtags in the Ushahidi map
- recruiting a "popular person" to endorse the software

Below is a collection of supporting answers for the aforementioned points, grouped by Formative interview questions. Most answers are paraphrased from the original responses.

How effective are the strategies used in the software at conveying environmental information?

- More explanation is needed of what is actually shown.
- We need legends/keys.
- I want to see things pop up when you hover instead of clicking on markers
- Perhaps recoloring map layers would show features more intuitively; i.e. vegetation would be colored green instead of red.

Are you able to make sense of environmental features over space?

- Generally, yes, but it's hard to differentiate features on the Risk Index map. It all looks red.
- The data viewer loads slowly (Indonesia-based response).

Are you able to make sense of environmental features in a particular space?

- No, I can't see the layers when zooming into Jakarta.
- It's not clear which layer is on top when turned on
- I don't understand the black parts of the risk map.

In what ways does the visualization help you understand flood-related risks?

- A slider for pixels would allow people to focus on specific levels within a risk map/ layer.
- It needs more explanation so that people know what we're talking about - it took me a while to see where the risk was.

In OpenIR's data viewer, do techniques like color variations and scales, time scales, and other "orienting" devices support or hinder your understanding of OpenIR's maps?

- It would be useful to add a layer with street names.
- It would be useful to change opacity of layers or be able to locate specific points.

In the Ushahidi environment, how do techniques like eyewitness report submission, report visualization, and other "crowd map" devices support or hinder your understanding of OpenIR's maps?

- They're absolutely vital. The big issue is the integrity of the crowd-mapping reports. Especially the responsible use of social media. And have realtime data, as opposed to the current Jakarta Globe slowness.
- Yes, but you should redo the Twitter keywords so you're not getting useless data: "flood," or in "banjir" Bahasa Indonesia; "Small flood" or "genangan" in Bahasa Indonesia. @TMCPoldaMetro is the biggest information source about traffic/floods in Jakarta. It's supported by police and corporations
- In Jakarta, there are neighborhood boundary lines, and people know their neighborhood number. People will reference locations by landmarks and street names. Much easier to communicate about floods using this information.
- Perhaps hand this deployment to local media groups to monitor.

How useful is the software ?

- The data viewer is more valuable for long-term infrastructure planning, and the data isn't going to change much over time. As far as current applicability, the data viewer may be best bet, until crowd-mapping begins to take off. But the applicability depends on having disaster data included, especially like storm surge info. The Tweets have great potential, but need to be reliable.
- For me it's very useful. For other researchers, it's very useful as well. But for general people, you'll need a lot of information, to guide the user, especially with the heights of water level.

Are you able to make meaningful interpretations of what you are seeing?

- It is difficult for us to read and interpret.
- If street names are visible, it'll be easier to interpret the information.

What is the role of prior experience in your ability to make sense of what you are using?

- Make the learning curve smaller. When people for the first time see the map, most Indonesian internet users are used to accessing Google Maps, so they will expect to see true-color images. They will then learn that this is not Google Maps, and they will need to know what it is the meaning of the different colors, where is the info coming from, how to interpret it, how to make it valuable for themselves, and they'll try to find the info.
- There should be text info for the first-time user so they can understand what kind of info they can get from the map.
- Make a layer label so that people know what they are looking at immediately

Would you use OpenIR going forward? How and why?

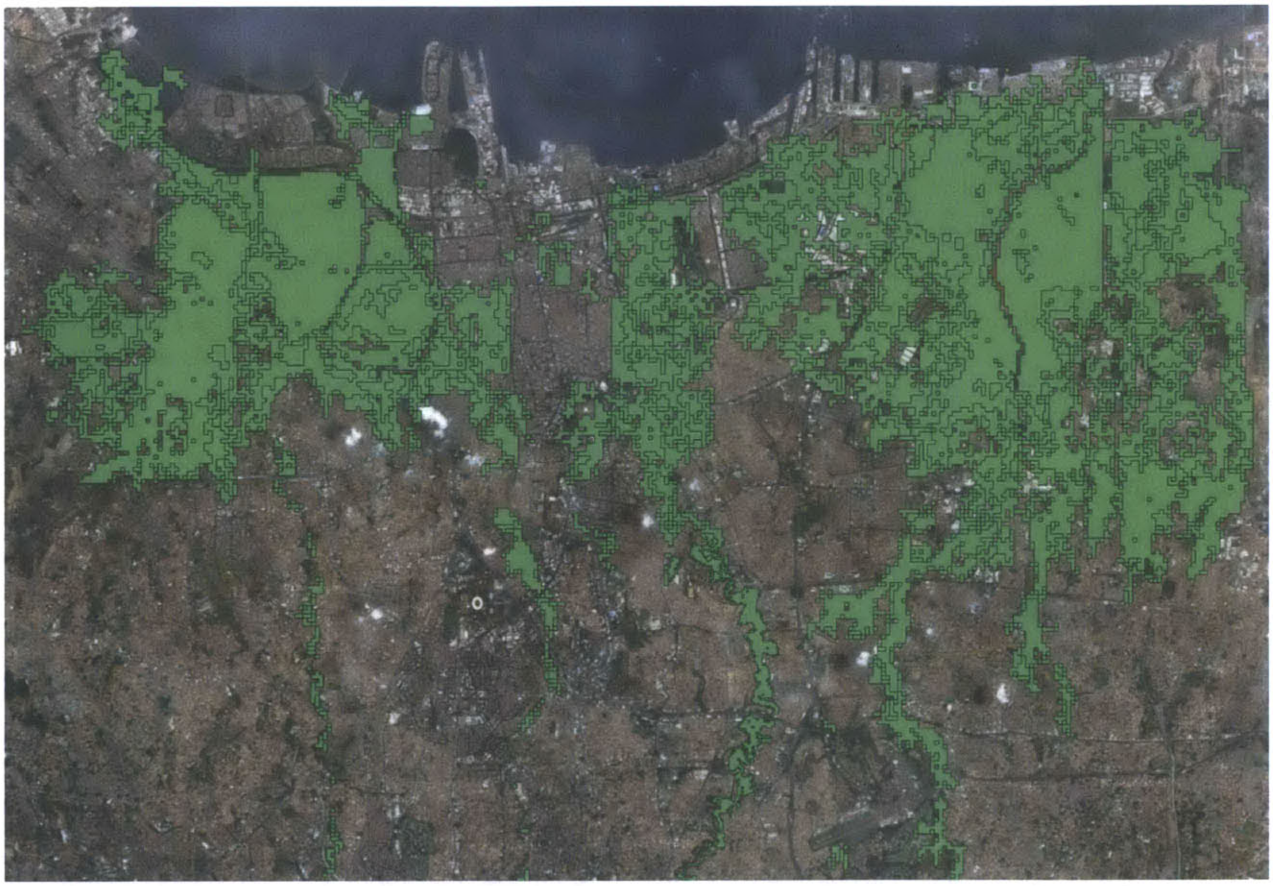
Yes, just make it user friendly. Especially for disasters and deforestation.

What additional support do you need for OpenIR to be useful to you?

- Make a mobile app.
- Get a popular person to be seen using the software
- Authorize people to put info on the map immediately.

Our team was not able to address all responses in just a few design+code sessions, but it addressed some of the most suggested key points in its Summative Data Viewers. The responses from both the Formative and Summative evaluations will processed more extensively, when the OpenIR effort continues beyond the scope of this thesis.

Chapter 10: Data Validation, Jakarta 2007 Flood
November 2012 to December 2012



Screenshot of InaSAFE Jakarta 2007 flood extents, in green. This image is overlaid on a Google Satellite image to show the flood's proximity to Jakarta's northern coast line.

AIFDR (Australia-Indonesia Facility for Disaster Reduction) maintains a plugin called InaSAFE for the open-source platform QGIS (Quantum GIS). This plugin offers a number of datasets and risk evaluation processes, including a map of Jakarta 2007 flood extents. Jakarta is located below sea level, on the coast, and between several mountains,[71] so it tends to flood every 3-5 years, during the height of its rainy season in January/February.[72]

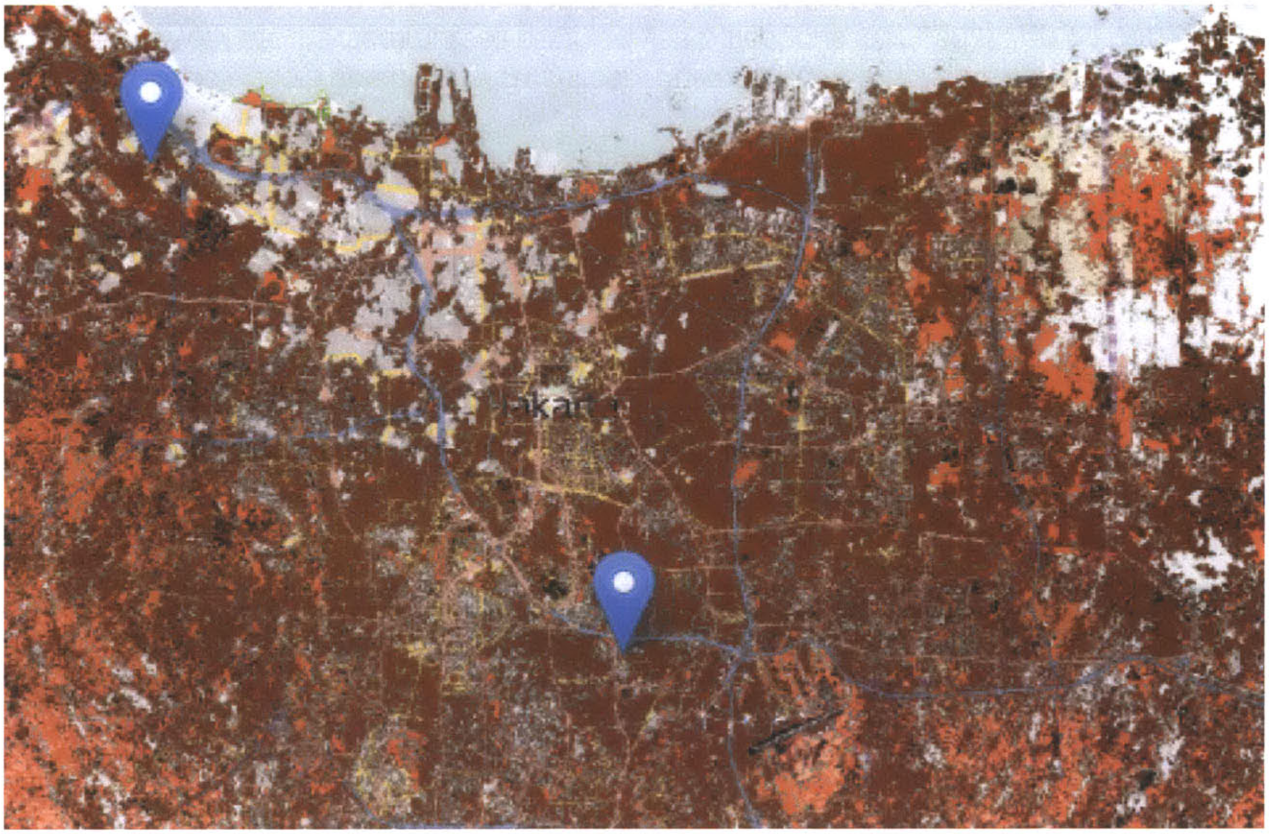
---

[71] Berg, Nate. "Jakarta is Sinking Itself Into the Ocean." 26 April, 2012. http://www.theatlanticcities.com/neighborhoods/2012/04/jakartas-sinking-itself-ocean/1857/

[72] Private conversation with Emmanuel Migo, January 2013.

The rapid development of its built environment has also contributed to flood extents.[73] Its 2007 flood was one of the most extensive in recent decades.[74]

After OpenIR team member Barry Beagen found this plugin, undergraduate researcher Stephanie New and I conducted a data validation process with this map, comparing it with OpenIR's Jakarta Risk Index map. We documented this process so that it can be replicated with other OpenIR Risk Indices and other observed data, including that for Hurricane Sandy in New York and 2013 flooding in Jakarta.



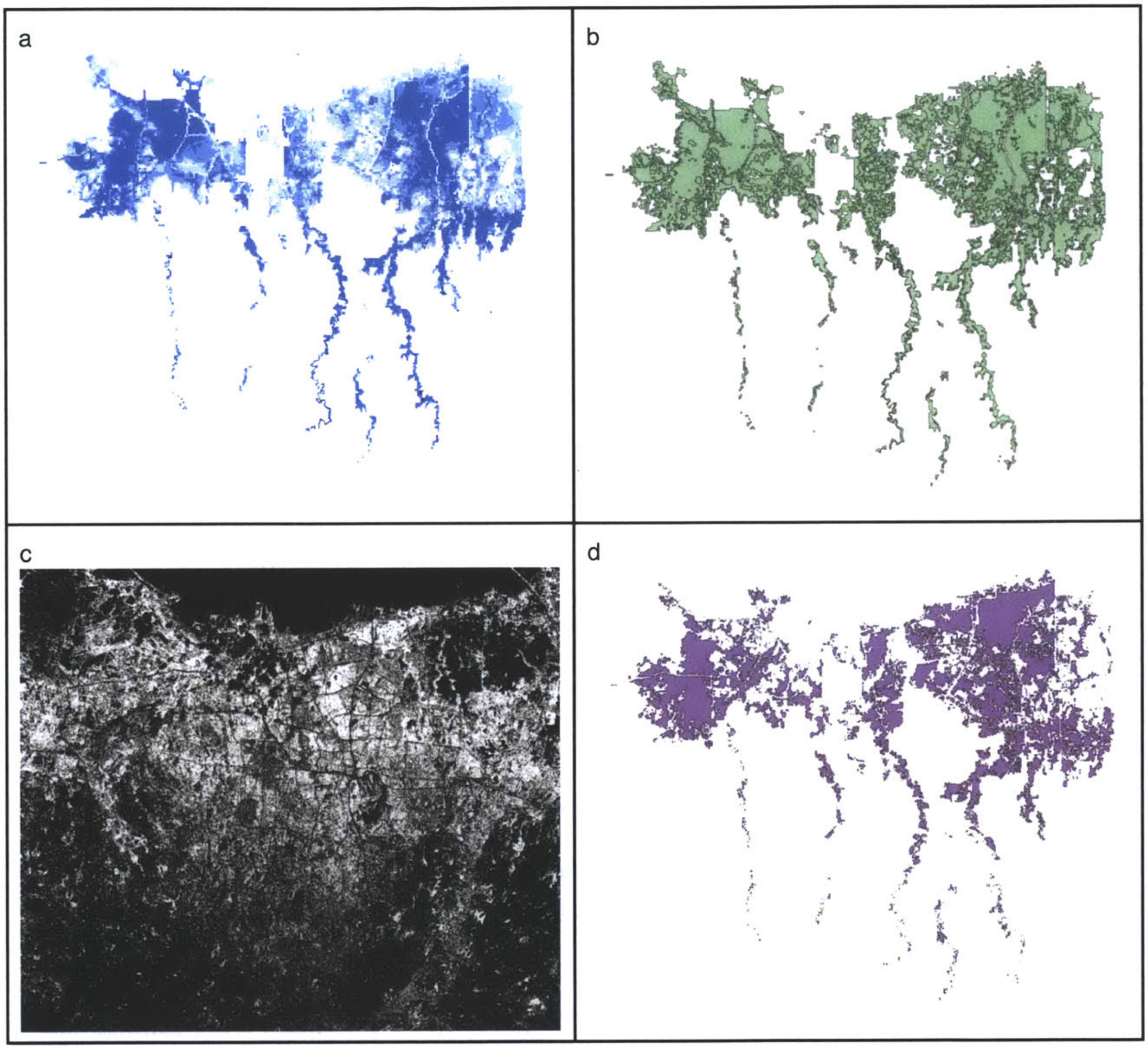
Screenshot of OpenIR Jakarta Risk Index map, in the northern coastal area flooded in 2007. The blue markers show points of interest in the OpenIR data viewer.

---

[73] Sarah Goodyear. "Flooding in Jakarta: A CIty Swamped by Its Own Success." The Atlantic Cities, 17 January 2013. http://www.theatlanticcities.com/jobs-and-economy/2013/01/flooding-jakarta-city-swamped-its-own-success/4426/

[74] "Jakarta floods death toll rises." BBC News, 4 February 2007. http://news.bbc.co.uk/2/hi/asia-pacific/6328873.stm

Validation Process



Screenshots of some intermediate imagery: a) the InaSAFE raster image, which contains depth information; b) InaSAFE imagery converted to vector format; c) the raw image for the OpenIR risk index level, in this case urban low-lying; d) the urban low-lying layer, clipped by the InaSAFE layer.

The general validation steps, mostly conducted in QGIS, consist of

- Converting, if needed, the observed flood extent image from a raster to vector image.
- Converting each Risk Index level (vegetation low-lying, vegetation high-lying, urban low-lying, urban high-lying) from a raster to vector image.

- Checking that the observed data and the Risk Index map are correctly geolocated and in the same coordinate system.
- Using the flood extent image as a clipping mask for the risk index image.
- Calculating the area of vegetation low-lying, vegetation high-lying, urban low-lying, and urban high-lying in the resulting image.
- Exporting these areas to spreadsheets.
- Calculating percentages of these areas (from the total area) in the resulting image.

In this particular validation process, the InaSAFE data contains information on flood depth (see figure below). However, for the purposes of this initial validation, we decided to discard this information, so as to more easily convert the data to vector form, and so as to restrict our process to a one-to-one comparison between flood extent and risk index levels.

The steps for the Jakarta 2007 validation process can be found in Appendix B.

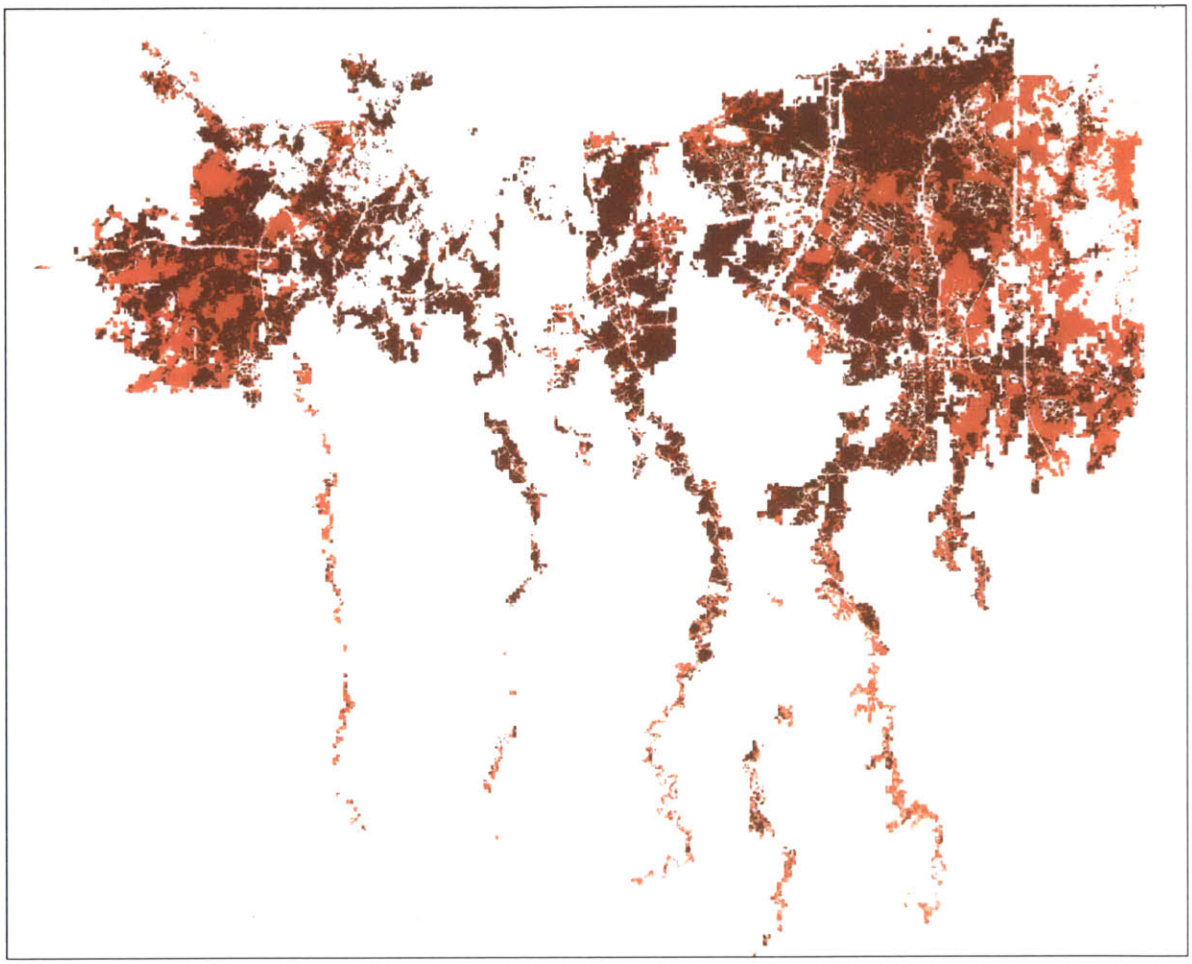
Validation Results

The final image, in which the observed 2007 data is a clipping mask for the risk index map, displays mostly dark red areas, indicating that the areas evaluated by OpenIR to be of highest vulnerability (urban low-lying) have the most overlap with 2007 flood extents. The resulting image displays very little light pink areas, indicating that areas evaluated by OpenIR to be a lowest vulnerability (vegetation low-lying) were not flooded in 2007.

Percent area calculations quantitatively confirm these visual observations: the highest index level (#4, urban low-lying) overlaps with 66.64% of the 2007 flood area, while the lowest index level (#1, vegetation high-lying) overlaps with only 0.47% of the 2007 flooded area.

OpenIR Jakarta Risk Index Map, Areas Flooded in 2007			
Index #	Feature	Color	% Overlap with 2007 Data
4	urban low-lying	dark red	66.64
3	urban high-lying	red	4.52
2	vegetation low-lying	salmon	28.38
1	vegetation high-lying	light pink	0.47

However, the middle indices, #2 and #3, are more problematic. At 4.52%, index level #3 is much lower than index level #3, which is 28.38%. This indicates that areas of low-lying vegetation was more extensively flooded than high-lying urban areas.



Screenshot of final result. The 2007 flood extent was used as a clipping mask for OpenIR's Jakarta Risk Index map.

In some ways, this result is quite intuitive: low-lying areas were more extensively flooded than high-lying areas; about 95% of the flood affected a low-lying area. But this process does highlight a problem in ranking environmental features in terms of risk. Perhaps it would make more sense for OpenIR's data viewers to just to show floodplains, that is, low-lying areas. On the other hand, it is useful to know how low-lying areas intersect with vegetation and urban surfaces. Perhaps, instead of numerically ranking feature combinations, it would make more sense to label what they are: urban low-lying, urban high-lying, vegetation, low-lying, and vegetation high-lying. This would help distinguish our map as an *index* and not a *predictive model*[75].

Professor Cesar Hidalgo also suggested that the four index colors are too similar; it would be easier on the eye if they were four different hues, instead of four shades of red.

Since conducting this validation, we discussed the process with Kristy Van Putten of AIFDR and Jesse Rozelle of FEMA, being careful to point out that the OpenIR Risk Index Map is an index that covers most land area, as opposed to a predictive model that shows more limited coastal areas. Both organizations are interested to see how the OpenIR Risk Index can be used as a general map to help affected populations consider their habitats, as opposed to a prescription for evacuation or direct emergency management.

These considerations return the discussion to issues of design and user interface, which we continued to explore in implementing our Summative Data Viewers.

---

[75] Scott MacKenzie: "Modeling Interaction:Descriptive and Predictive Techniques." York University, 2007. http://www.cse.yorku.ca/course_archive/2007-08/F/6329/03-ModelingInteraction.pdf

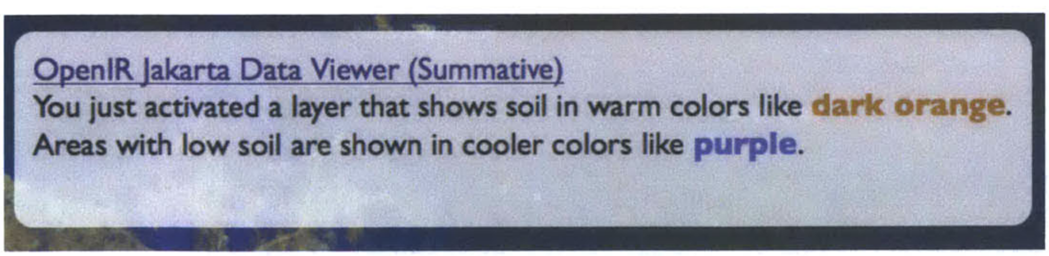
Chapter 11: Implementation Phase 4
January 2013 to February 2013

Implementation Phase 4 was short and intense, as it overlapped with the OpenIR team's travel to Indonesia. As a result, January 2013 was not just a time of technical implementation, but also of community development, surveys and interviews, outdoor fieldwork, and cultural observation. Only a few technical features were implemented in Phase 4; the fuller Indonesia experience is discussed in the next chapter.

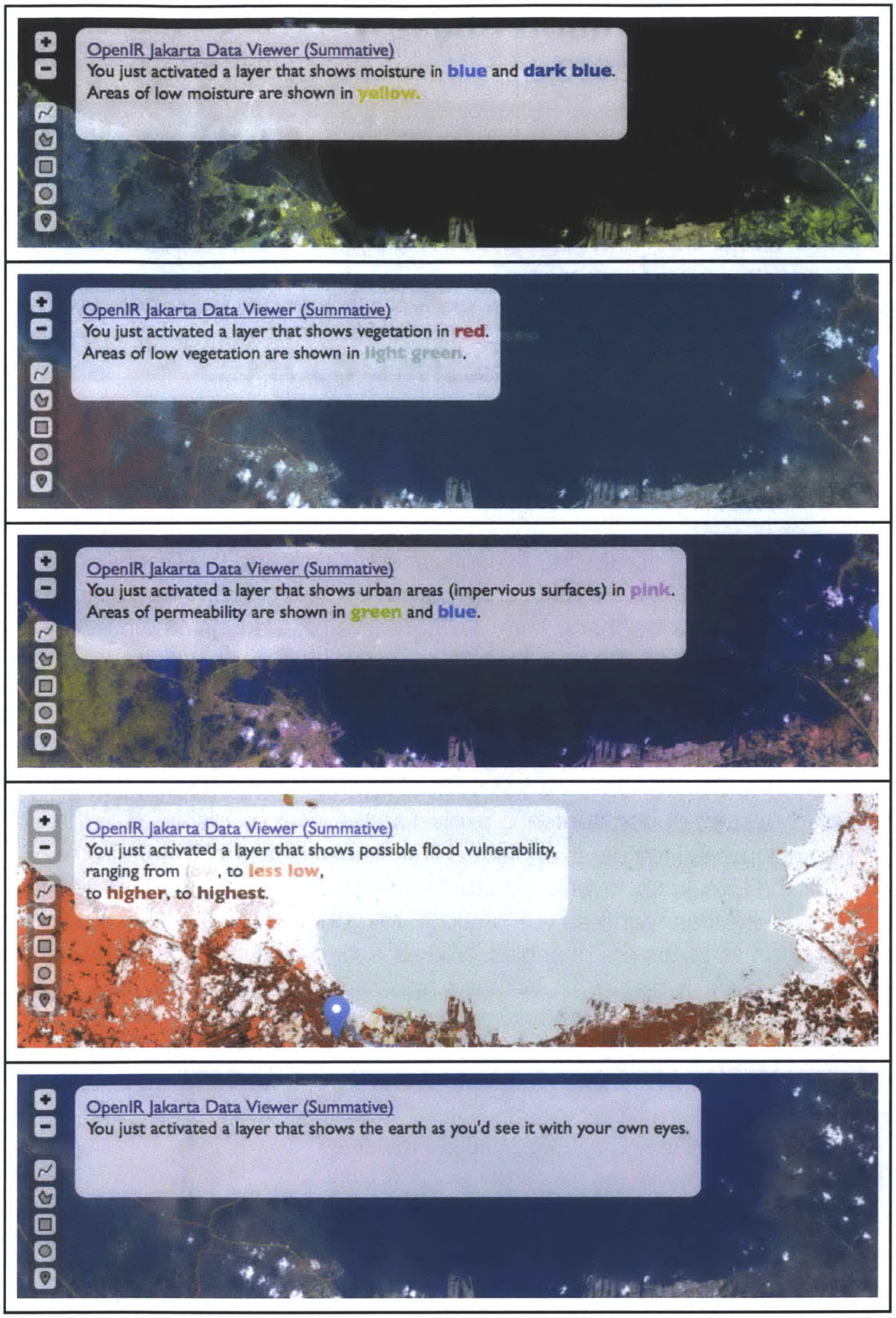
Summative Data Viewers

After Formative Evaluations were completed in early January 2013 and the OpenIR team traveled to Jakarta on January 8-10, 2013, several days were spent modifying versions of the Formative Data Viewers with feedback from the Formative Evaluations (see previous chapter). The result was a set of Summative Data Viewers for Jakarta and Java/ Kalimantan.

The most common Formative Evaluation suggestion was to include a legend or key explaining the colors in the OpenIR Data Viewer layers. For this, I implemented, using Javascript and JQuery, a color-coded sentence explaining the layer on display. For the IR layers, the salient environmental feature is explained first, followed by the feature in the contrasting color. For the Risk Index layers, the four index colors are explained, while the True Color layers is explained as "earth as you'd see it with your own eyes." These color-coded sentences change as layers are changed; when a layer deactivated, the statement "You just de-activated a layer" is displayed. With this approach, light colored words, tended to be less legible; the implementation was clearly provisional.

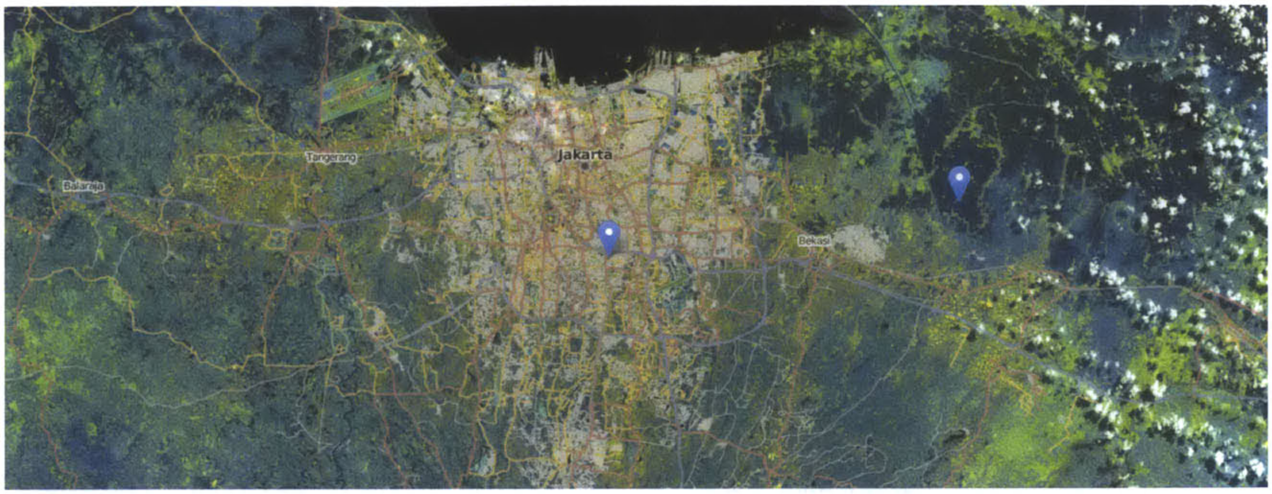


Screenshot of color-coded sentence explaining the "Soil" layer.

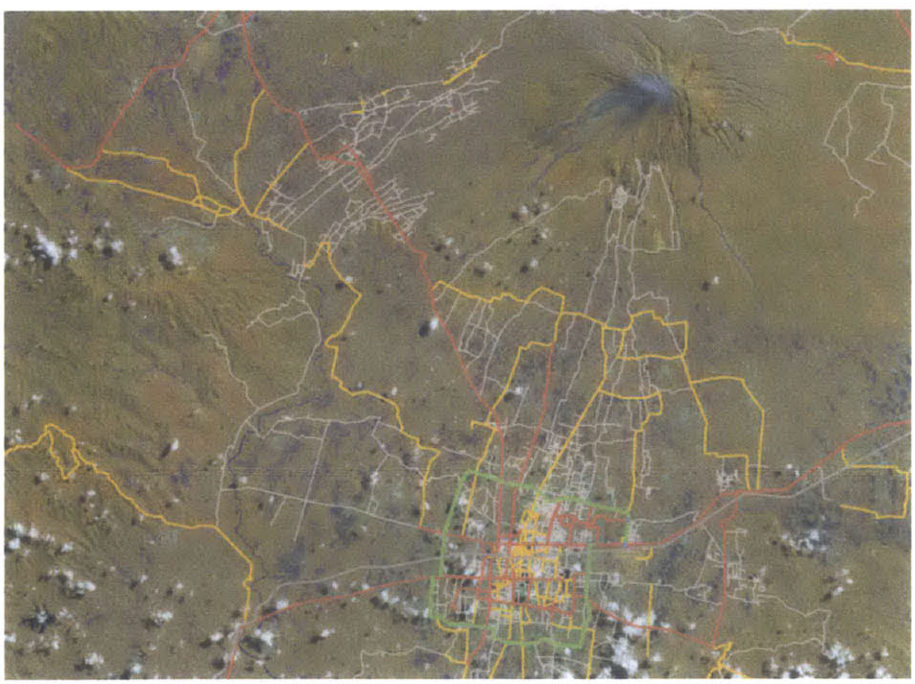


Screenshots of explanatory statements for other layers. This approach is only provisional, especially since colors like light pink, light green, and yellow were hard to read.

Another common Formative Evaluation suggestion was to simultaneously show streets with IR data, so that users could better see their homes' proximity to environmental features and risks. For this feature, I applied a modified OpenStreetsMap layer with only streets and city names, available from Cloudmade.[76] This layer did not have a solid background, making it suitable for overlaying on raster maps. A slider on the Data Viewer's upper-right corner allows users to control the opacity of the street overlays.



OpenIR Jakarta Data Viewer screenshot of streets overlaid on IR Moisture layer. The street opacity is 50%.



OpenIR Java/Kalimantan Data Viewer screenshot of streets overlaid on IR soil layer. Street opacity is 100%, making it easier to see the proximity of the city Yogyakarta (lower right) to the Merapi volcano's burn scars (upper right).

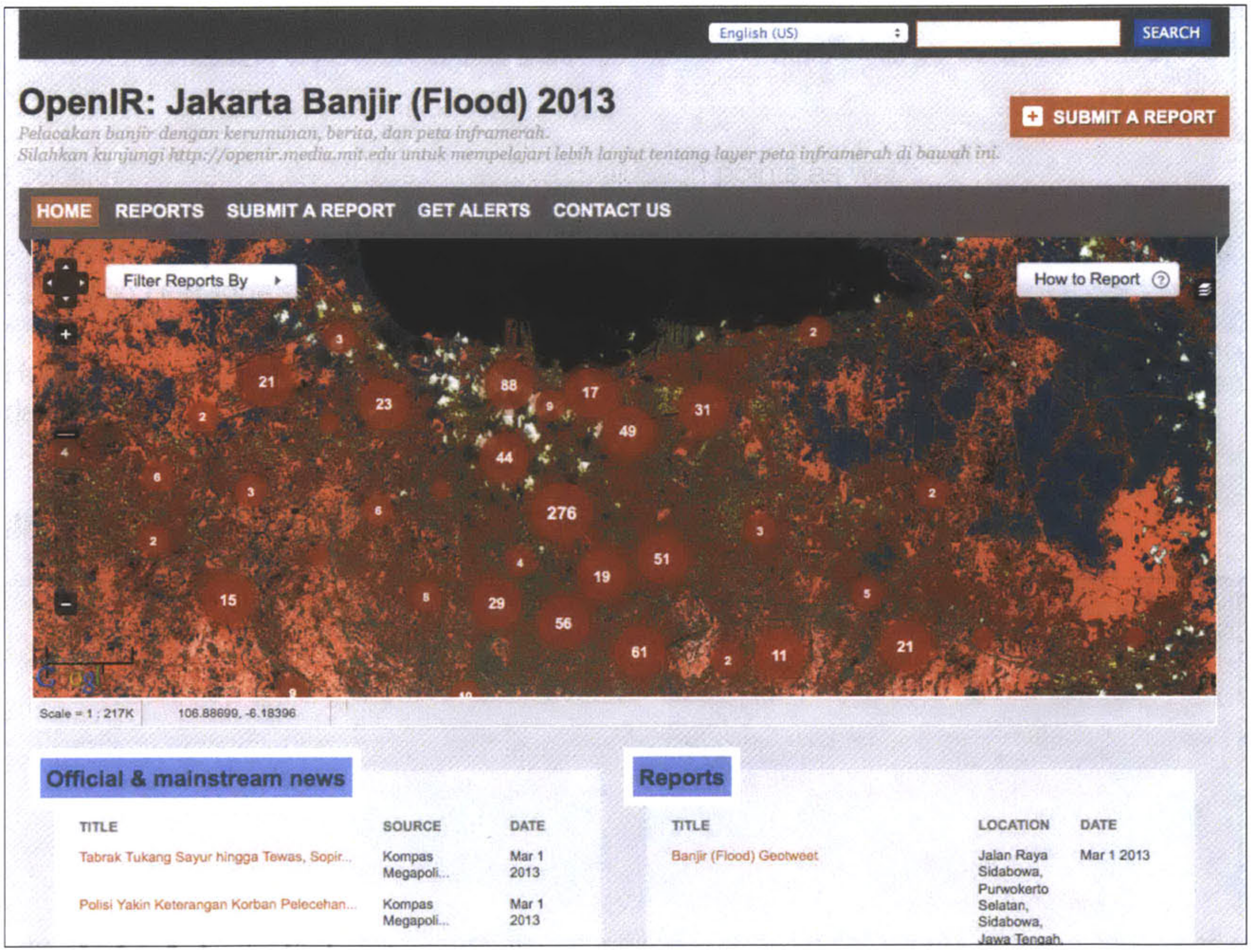
---

76 Cloudmade Maps Editor: http://maps.cloudmade.com/editor

I considered addressing these Formative Evaluation points as well:
- processing additional zoom levels
- implementing hover-based instead of click-based explanations of environmental features
- allowing Risk Index levels to be toggled separately

However, due to limited implementation time, I decided to leave these features for a later date.

Jakarta Ushahidi Deployment



Screenshot of OpenIR's Ushahidi Deployment for Jakarta's January 2013 flooding.

On January 16-18, 2013, heavy rains caused flooding that paralyzed much of Jakarta. With many roads, highways, and buildings flooded up to several meters deep, schools, businesses, and government offices were forced to close for several days. The OpenIR team was negatively impacted, as explained in the next chapter. One positive outcome is that, from my unaffected apartment, I was able to modify the OpenIR/Ushahidi Jakarta deployment, built as an experiment in November 2012, to a point in which it could be used, edited, and understood by Jakarta citizens affected by the crisis. These modifications included

- narrowing the deployment's geo-Tweet hashtags from "Jakarta" to "banjir" ("flood" in Indonesian), "jakartabanjir2013," and similar terms. This better filtered the kinds of messages that were visualized on the map. Because of our team's human resource limitations, most of the reports to the Jakarta Banjir 2013 map were scraped from Twitter, not actively logged via the Ushahidi platform.

- including newsfeeds in the Indonesian language.

- activating SMS notifications for interested subscribers.

- translating the deployment's title and subtitle to the Indonesian language.

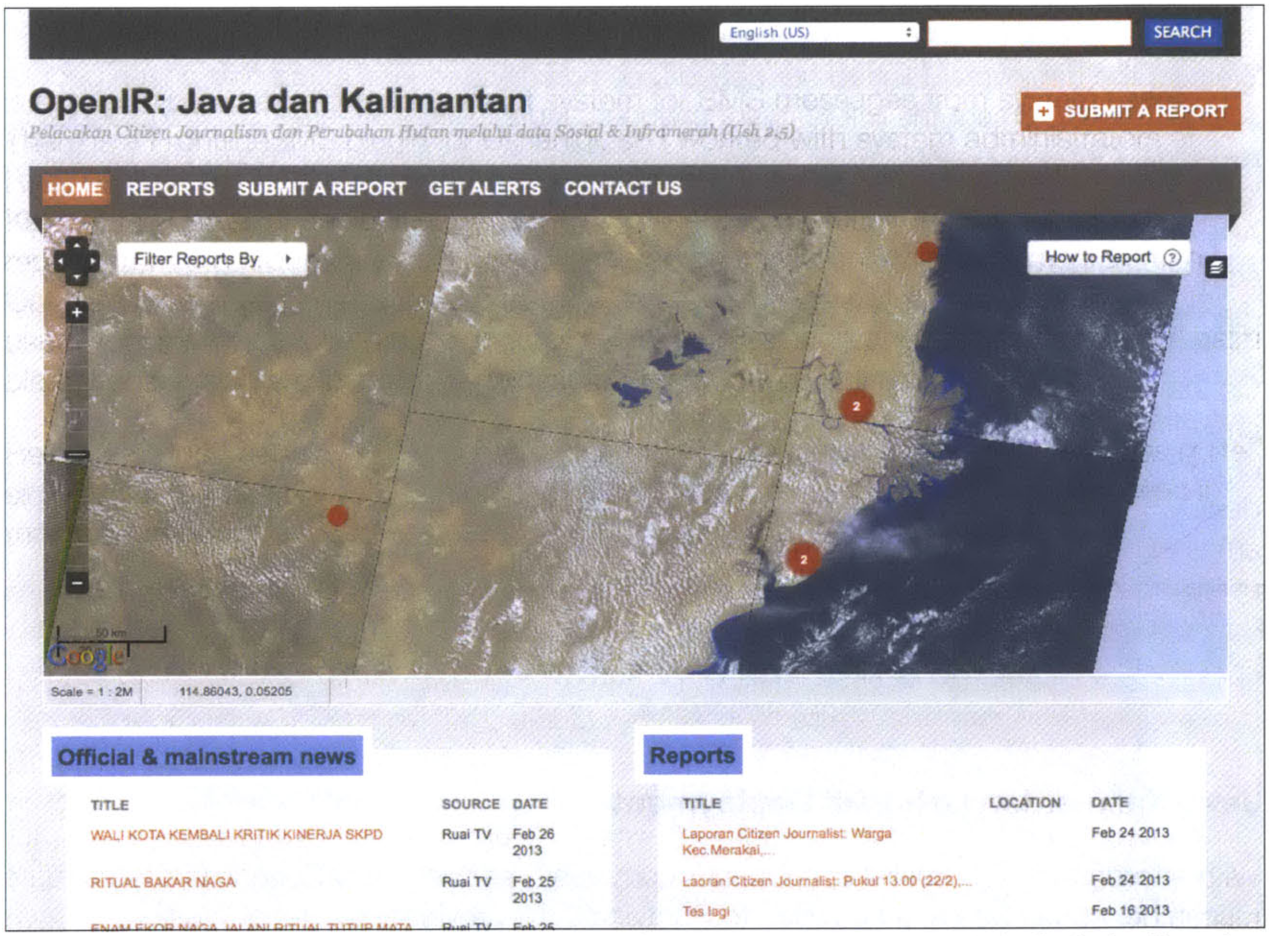
We sent the deployment link to our network community partners, about two dozen people in all. They in turn sent the link to their networks, and a few volunteers started to suggest revisions, report bugs, and enter their flood information into the deployment. The OpenIR team at this time was also sent observed January 16th, 17th, and 18th flood data from the Australia Indonesia Facility for Disaster Reduction, which it started to visualize in its standalone Jakarta Data Viewer. It was a gratifying moment for our tiny team to have reached the point where our work was being used in an actual crisis.

On the other hand, with no crisis responders actively working on the deployment, the crowd data collection was primarily from Twitter. This data was not verified, categorized, or curated. As a result, while there are more than 1500 crowd data points on the deployment map, many of which overlap with areas indicated by OpenIR's Risk Index as areas of high vulnerability, it will take much further work to export and clean the data for useful analysis with the OpenIR layers.

Java/Kalimantan Ushahidi Deployment

With environmental journalist and OpenIR community partner Harry Surjadi, undergraduate researcher Juhee Bae and I traveled to Pontianak, the provincial capital of West Kalimantan, and Palankaraya, the provincial capital of Central Kalimantan. In Pontianak, Harry worked with Ruai TV, a local network covering indigenous issues, particularly those of the majority Dayak tribe, to set up a citizen journalism system that they call "RuaiSMS." In Palankaraya, Harry set up similar systems with the groups Sekber REDD+ (Sekretariat Bersama, Reducing Emissions from Deforestation and Forest Degradation) and AMAN (Alliance of the Indigenous Peoples of the Archipelago) to set up similar systems. All of these systems were run with FrontlineSMS, an open-source software in which a computer is set up to be a content management system for SMS messages from system users. There is a FrontlineSMS plugin for Ushahidi, so I worked with system administrators at Ruai TV, Sekber REDD+, and AMAN to activate this plugin in both the OpenIR Java/Kalimantan Ushahidi deployment and the FrontlineSMS systems. As a result, citizen reports from the groups are now being visualized on the OpenIR Java/Kalimantan map. Our hope is that as both social and satellite data accumulates in this map, it will be a useful tool for indigenous groups in conflict with deforestation policies of corporate oil palm plantations, which are increasingly widespread throughout Kalimantan.

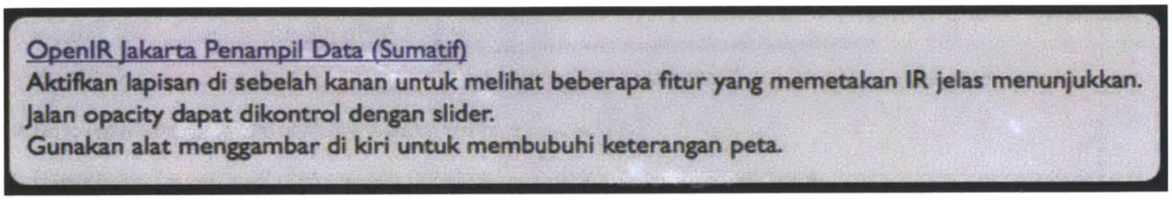
Being well documented by the Ushahidi community, the technology initialization was quite simple, but the deciding on a reporting protocol was more complicated, as discussed in the next chapter.



Screenshot from OpenIR Ushahdi deployment for Java/Kalimantan. The map console is displaying Central and East Kalimantan, with a few initial citizen journalist reports.

Translations

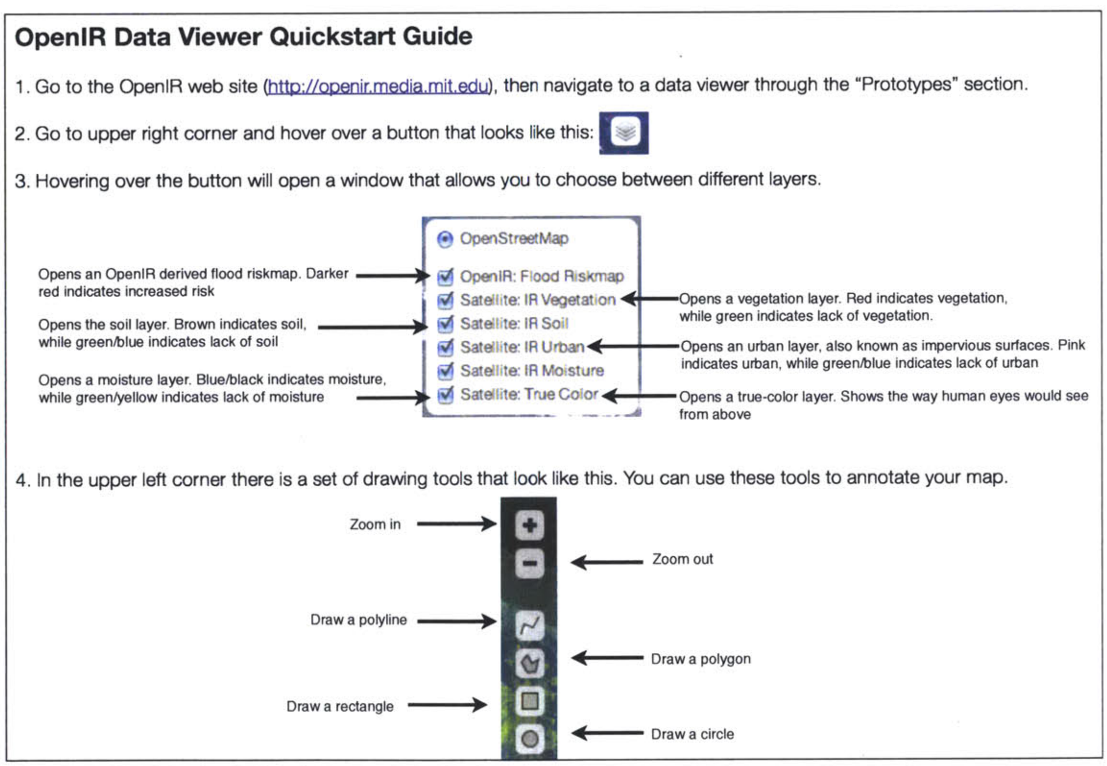
In many of our meetings, workshops, and presentations throughout Indonesia, our community partners pointed out that the OpenIR Data Viewers were only in English. This kept many people from being able to thoroughly evaluate the Data Viewers. In response, undergraduate researcher Juhee Bae and used Google Translate and the assistance of Professor Trias Aditya to translate OpenIR's Data Viewers and Summative Evaluation surveys to the Indonesian language.



Screenshot of OpenIR Jakarta Data Viewer title bar, in the Indonesian language.

Quickstart User Guides

Before our Indonesia trip, undergraduate researcher Juhee Bae and I developed a "quickstart" user guide to help new users quickly understand how to use OpenIR's data viewers. We printed several dozen guides, especially in anticipation of internet connectivity issues in more remote parts of Indonesia.



OpenIR Data Viewer Quickstart Guide

1. Go to the OpenIR web site (http://openir.media.mit.edu), then navigate to a data viewer through the "Prototypes" section.
2. Go to upper right corner and hover over a button that looks like this:
3. Hovering over the button will open a window that allows you to choose between different layers.

4. In the upper left corner there is a set of drawing tools that look like this. You can use these tools to annotate your map.

OpenIR "Quickstart" Guide.

Chapter 12: Indonesia Fieldwork Report

Indonesia emerged as an initial case study area for OpenIR because it is 1) seismically active, agriculturally productive, and tropical, thus prone to rapid environmental change; 2) economically developing and thus open to experimentation in ecological information communication technologies (ICTs). Indonesia was the first host to some major data innovation initiatives, including the Humanitarian OpenStreetMap Team (HOT) and UN Global Pulse, both of which became OpenIR's first community partners.

In January 2013, with funding in large part from MIT's Public Service Center, OpenIR's Usability Researcher Juhee Bae traveled with me for three weeks in Indonesia. OpenIR collaborator Barry (Wirajaya) Beagen, a Jakarta native, joined us during our first week, while collaborator Ilias Koen teleconferenced in to our meetings throughout the entire trip.

Because infrared satellite data is used for so many environmental purposes, our primary goal was to present and evaluate OpenIR with a range of stakeholders focused on a range of environmental issues. Based on these presentations, we evaluated which groups and issues seem the most promising for developing future OpenIR applications.

OpenIR's cast of community partners has grown and re-balanced since its inception, but Indonesia has proven to be an amazing place to conduct initial usability and data validation studies, in large part to these open-minded communities. In addition to HOT and UN Global Pulse, these communities include the Indonesian Ministry of Communication, Ruai TV, AMAN (Aliansi Masyarakat Adat Nusantara: Alliance of Indigenous Peoples of the Archipelago), Sekber REDD+ (Sekretariat Berasama / Reducing Emissions from Deforestation and Forest Degradation), Universitas Gadjah Mada, Combine Research Institute, and AirPutih.

Three individuals were particularly instrumental to OpenIR's progress in Indonesia: Fadhilah Mathar of the Indonesian Ministry of Communication, who helped my team settle into Jakarta; Kate Chapman of HOT, who connected us to additional partners and kindly hosted us in the final days of our trip; and especially Harry Surjadi, who introduced OpenIR to his Information Broker system of environmental citizen journalists throughout Kalimantan (formerly Borneo). Thanks to Harry's foresight, OpenIR will likely play a key role in future Information Broker maps.

Timeline

<u>Week 1</u> focused on flooding in Jakarta, where Juhee and I were joined by our teammate Barry, whose family lives in the northern district of Pluit. We met representatives from

groups including HOT and UN Global Pulse, Rujak.org, Indonesia's Ministry of Communication, Australia-Indonesia Facility for Disaster Reduction (AIFDR), Tempo News, Rujak.org, UNDP, and UN Habitat.

Many of our initial Javanese (Jakarta) meetings took place over "java," in cafés. Left: A meeting with Harry Surjadi and HOT. Center and right: discussions with Rujak.org, a design and planning non-profit in Jakarta.

All of our Indonesia meetings roughly followed this agenda:
- Opening remarks by host
- OpenIR overview presentation by me
- Demonstration of OpenIR data viewers by Juhee, sometimes assisted by me
- Discussion of OpenIR for community development, moderated by Barry (later Juhee)
- Q&A, moderated by me
- Structured evaluation using "talk-aloud" protocols, moderated by Juhee
- Final remarks by host

Left: Arlene and Juhee meet with Tanti Liesman, Coordinator of UN Global Pulse's new Jakarta Pulse Lab. Right: Elvina Anita and Fadhilah Mathar bookend the OpenIR team after a workshop at the Indonesian Ministry of Communication's ICT Training Center.

January is the height of rainy season in Jakarta. On January 16, heavy rains caused severe flooding, displacing thousands of people and trapping Barry in his home near the Jakarta coastline. After his home was flooded, Barry was briefly hospitalized for a bacterial stomach infection. It was a blow to the project, as Barry had to withdraw from OpenIR's trips to Kalimantan (see Week 2). On the other hand, the floods gave OpenIR a unique opportunity to transform its Jakarta data viewer, flood index map, and crowd map (Ushahidi) deployment, all developed in 2012, into active tools for evaluating the floods. After we presented OpenIR to Tempo.co, one of Indonesia's largest news outlets, they launched a simpler (uncategorized) crowd map of flood-related geo-tweets.[77] Kristy Van Putten of AIFDR sent our team flood level data from January 16; after returning from Indonesia, I started working with undergraduate researcher Wendy Cheang to validate OpenIR's Risk Index Map against this new data.

Left: Juhee and I made it through flooded streets to Tempo.co, one of Indonesia's largest news outlets. Center: In the lobby, TV news covered flooding. Right: Juhee demonstrates OpenIR to Tempo.co staff.

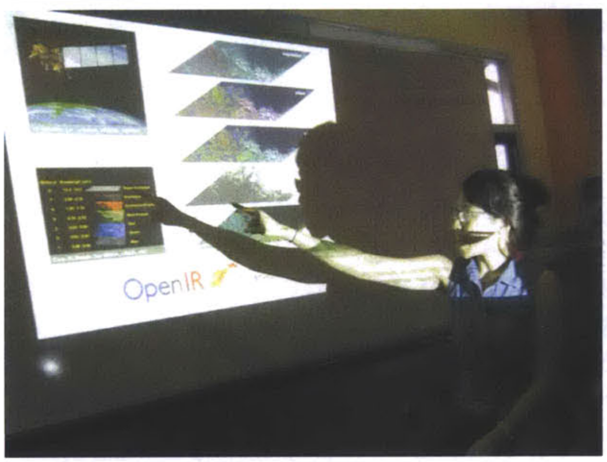
Week 2 focused on deforestation in Pontianak, the capital of West Kalimantan province, and Palankaraya, the capital of Central Kalimantan province. Guided by recent Knight Fellow Harry Surjadi, I connected his citizen journalism systems to OpenIR's data viewers. Surjadi, working with indigenous advocacy groups in West and Central Kalimantan, developed an expanded approach to citizen journalism called "Information Broker" systems,[78] in which locals are trained to use SMS to not only report, but to directly communicate with government officers on important municipal issues. Harry Surjadi flew with Juhee and me to Pontianak, where he initiated, facilitated, and translated a lively meeting at the headquarters of Ruai TV, a network focusing on indigenous peoples' issues throughout the province, particularly those of the Dayak, Kalimantan's predominant ethnic group.

Both the West and Central Kalimantan trips were two-day sessions, with the first day devoted to roundtable discussions with about a dozen environmental stakeholders. These stakeholders included our hosts, as well as local representatives from BNBP (Indonesia's emergency management agency), World Wildlife Fund, UN-ORCID, USAID, Kalimantan provincial government, and other organizations.

[77] The map is now down, but a related article is at http://en.tempo.co/read/news/2013/01/16/058454810/Banjir-Rendam-Puluhan-Sekolah-di-Serang

[78] Harry Surjadi: "Information Broker: Communicating and Monitoring Climate Change." http://harrysurjadi.wordpress.com/2012/09/19/information-broker-communicating-and-monitoring-climate-change/

Left: Environmental stakeholder meeting at Ruai TV. Right: Ruai TV interviews Juhee and me.



Left: I present OpenIR at Ruai TV's environmental stakeholder meeting. Right: Ruai TV staff after the meeting.

On the second day of these sessions, Juhee contacted, interviewed, and surveyed environmental stakeholders,while I worked to set up an infrared mapping system with Ruai TV which has a citizen journalism system established under Harry's previous fellowships. Over time, we think that citizen journalism reports, visualized on an infrared map, may show both locational and environmental patterns in aggregate. The technical backbone of our hosts' citizen journalism systems is FrontlineSMS, an open-source software widely used to manage texting for social change, as in the case of citizen journalism. Ushahidi has a FrontlineSMS plugin which the OpenIR had tested before the Indonesia trip, so the second-day sessions consisted of connecting the host organization's FrontlineSMS system to OpenIR's Ushahidi-Kalimantan deployment,[79] and planning a protocol for OpenIR to map received SMS messages. The technical part was straightforward, but the protocol planning took some extended discussion, for a few reasons:

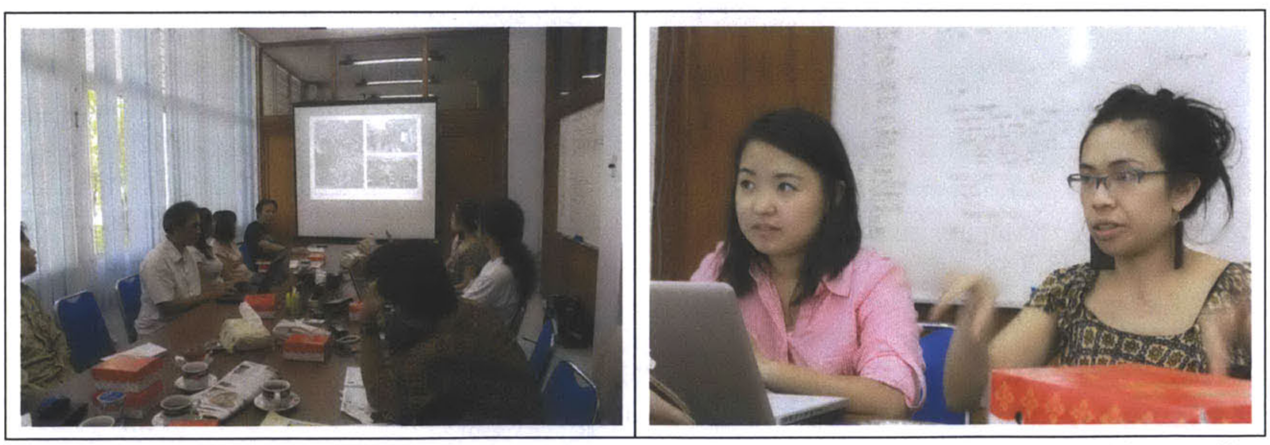
---

[79] OpenIR: Java dan Kalimantan. http://openir.media.mit.edu/Ushahidi/Java_Kalimantan/index.php/

- Most citizen journalists reporting to the Ruai TV, Sekber REDD+, and AMAN systems do not use smartphones; their phones lack GPS sensors and their SMS messages are not geo-located.
- These citizen journalists report in the Indonesian language, but OpenIR's maps are developed in the English language.
- Initial reports from citizen journalists are then edited by the host organization, which then blasts the edited SMS report to all subscribers of the relevant FrontlineSMS deployment.

In the end, we decided that editors would name the district location of the report before making the final blast. Every week, we would then use provincial district maps for West and Central Kalimantan to place the report on the OpenIR map. This approach will work, but its report placement will be low-resolution, as we could not find reliable provincial maps for sub-districts or streets. Also, as OpenIR's team members outside responsibilities ebb and flow, its abilities to manually place the reports every week may fluctuate. On the other hand, this manual system will better ensure that reports are properly placed; I asked SMS-mapping groups throughout the world (via Standby Task Force[80] and Ushahidi Skype chats[81]), about automated systems for placing non-geo-located SMS messages, and it seems that no geographically reliable automated system currently exists.

From Pontianak, Harry, Juhee, and I flew back to Jakarta and then immediately to Palankaraya, where we conducted similar meetings and workshops with Harry's colleagues at the headquarters of Sekber REDD+. Palankaraya, and its province of Central Kalimantan, is a kind of "poster child" for tropical forest restoration. The UN is investing much time and effort in this area, and many travelers to Palankaraya are environmental researchers.



Left: Harry Surjadi (near screen) facilitates an stakeholder meeting at the Sekber offices. Right: Juhee and I present OpenIR at this meeting.

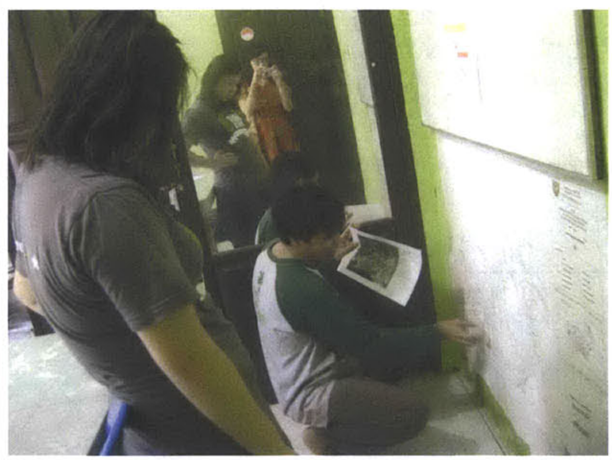
---

[80] The Standby Task Force: Our Model. http://blog.standbytaskforce.com/our-model/

[81] Technical Members - Community Wiki: Ushahidi. https://wiki.ushahidi.com/display/WIKI/Technical+Members

Left: Environmental Stakeholder meeting at Sekber REDD+. Right: Work session with Palankaraya citizen journalists at Sekber's offices, in which we install a OpenIR/Sekber map/sms connection and reporting protocol.

Juhee and I stayed one extra day in Palankaraya, allowing us to work with both Sekber REDD+ and AMAN, and to install an OpenIR map connection with both of their sms-based citizen journalism systems.



Work session with AMAN, in which we install a connection between OpenIR's map and AMAN's citizen journalism systems. Left: following custom, we work on the floor. Right: Dadut, AMAN Kalteng's leader, explains some indigenous boundary maps.

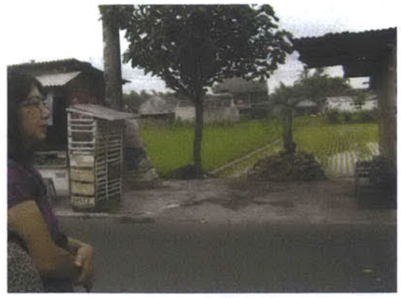
<u>Week 3</u> included a journey to Yogyakarta, where Universitas Gadjah Mada's (UGM) Department of Geodetic Engineering hosted Juhee and me for several nights. Yogyakarta is a college town, for it has the most universities of any Indonesian city. It is also a cultural town, for it was Indonesia's pre-colonial capital, and it is known for significant Buddhist and Hindu temples, as well as batik and wayang kulit (shadow puppet) workshops. Yogyakarta lies in close proximity to the volcanically active Mount Merapi, which most recently caused extensive damage and loss of life in its 2006 eruption. Remote sensing is

often used to analyze volcanic damage, and academic papers on Merapi risk mapping provided some of the inspiration for OpenIR's own approach to risk maps.

Presentation to Universitas Gadjah Mada's Department of Geodetic Engineering, in Yogyakarta.

While OpenIR is not including volcanic damage, which globally affects relatively few people, as part of its initial risk maps, it was illuminating to explore the lasting effect of such a dramatic environmental phenomenon. One of our hosts, UGM's Professor Heri Sutanta, drove Juhee and me to Mount Merapi and explained how pyroclastic flow incinerated homes, forests, and animals. Telling us that much of the volcano's tourist industry was run by affected residents, he signed us up for a motorcycle tour of the volcano's middle region. As an urban dweller who generally sees most of a region by foot or bike, it was rather thrilling to observe an environment through an open-air vehicle like a motorcycle. I had a similar feeling when we flew back to Jakarta; the route took us over some of the irrigated area that we discuss in OpenIR demos, and it was heartening to directly see some of these areas, many of which are uncharted in street maps.

Our final roundtable discussions were to UGM's Geodetic Engineering Department, the locally-founded NGO Combine Research Institute, and back in Jakarta, to our first community partner back, the Humanitarian OpenStreetMap Team. By now, Juhee and I had our presentation procedure down pat, and were able to collect much feedback from the usability sections of these presentations.



Final workshops in Indonesia. Left: OpenIR at the Combine Research Institute in Yogyakarta. Center: standing with Ade Tanesia near the rice farm outside Combine. Left: Humanitarian OpenStreetMap Team in Jakarta. Most of the team consists of recent geography graduates from local colleges.

Flying back to Jakarta from Yogyakarta, I saw land that we often highlight in OpenIR presentations. This land, just east of Jakarta, is unmapped in OpenStreetMaps and other road maps, and highlighted as strongly irrigated on IR moisture maps. It was fascinating to see this land from the air, and to realize that while IR maps can't capture its beauty, it can help to show how that land exists and is used by humans.

Turning Points

By the conclusion of our trip, both Juhee and I were quite satisfied with the range of stakeholders we met and the range of experiences we encountered. We had meaningful conversations and collected a lot of data, sometimes at the same time, and we more clearly identified a target user for OpenIR. That said, my most clear "A-ha!" moment during the trip was at the height of the Jakarta floods, on January 16-18, days on which most businesses and schools closed or loosened their attendance policies. While the floods complicated our meeting schedule, being stuck in our apartment during these days allowed me to polish up the Jakarta Ushahidi deployment that had been in R&D for months, born from a couple of Geeks without Bounds hackathons in early autumn 2012. Our goal back then was to build an OpenIR plugin for Ushahidi, and this allowed us to plug-in our data to an Ushahidi deployment. We also set up geo-Twitter scraping, so that by the time the Jakarta floods were in full force, we had overlaid more than 1500 crowd reports on to the OpenIR risk map. With this many reports, it was easy to see the similarity between the placement of the reports and the areas colored as "highest risk" (urban low-lying areas) in the map. It was exciting to bring this map from an R&D to an active tracker phase, and to feel that the map was playing a small role in flood response. The long period of gestation for this map paid off quantitatively when Ushahidi gave it a Deployment of the Week award.

This triumph is small, though—now that I have played a role in compiling both IR and social data for an environmental disaster, I can see a number of issues in its use for longer-term analysis and recovery. With the IR data, the OpenIR risk map marks urban high-lying areas as higher risk than vegetative low-lying areas, but ultimately all low-lying areas, both paved and planed, are more prone to flooding. With the social data, it is tricky to verify the accuracy and relevance of reports from so many formats (tweets, sms, emails, web forms), and it's currently difficult to export this data for further analysis. I've noticed these issues in maps by us and by others, and this is a major issue raised by Professor Trias Aditya, our primary host at UGM. Almost everyone in the web mapping community envisions better data interoperability and uniformity, so that data can more easily move from map to map. This will not only help users, but also researchers and analysts. As OpenIR continues beyond the scope of this thesis, I hope to better analyze how social and IR data inform each other and can be used together.

Chapter 13: Summative Evaluations

OpenIR's journey to Indonesia launched a period of Summative Evaluations, which consisted of "talk-aloud" sessions and surveys focused on user feedback of the OpenIR Summative Data Viewers. The talk-aloud sessions were segments of OpenIR's presentations throughout Indonesia, in which undergraduate researcher Juhee Bae demonstrated the OpenIR software, asked a series of structured evaluation questions to meeting attendants, collected attendant contact information, and sent attendants a follow-up survey to take after directly interacting with OpenIR's Data Viewers.

Summative Interviews and Talk-Aloud Sessions: Initial Findings

In Indonesia, we conducted nine presentations/meetings/workshops, and solicited feedback from 71 respondents:

UN Global Pulse: 1
John Taylor / UNDP: 1
Tempo.co: 7
Combine Research Institute: 6
Ruai TV (Pontianak environmental stakeholders): 14
Humanitarian OpenStreetMap Team: 9
REDD+ (Palankaraya environmental stakeholders): 18
Indonesian Ministry of ICT: ~8 (not all attendants provided contact information)
Universitas Gadjah Mada: ~7 (not all attendants provided contact information)

We are still analyzing feedback collected to date, including back-end analytics for the data viewers used in Indonesia. We are also still collecting new feedback from online surveys. But at this point, it is helpful to summarize the findings from our "talk-aloud" usability discussions.

Before we traveled to Indonesia, our community partner Kate Chapman cautioned that in many of her own workshops, mappers found infrared satellite images to be too "weird." I wasn't sure if OpenIR would gain traction through our presentations, but I was pleased to find that almost all of our audiences were enthusiastic about using IR data for their purposes, acknowledging the IR data could help fill information gaps left by streetmaps and true-color maps. From our discussions, the average "usefulness" score for OpenIR was about 9 out of a possible 10. I think part of this high score was due to our presentation, which I think explained OpenIR's underpinnings and functionalities quite thoroughly. Our discussions range from 2-4 hours in duration, which gave audiences much time to ask extensive questions about OpenIR and IR data in general.

OpenIR's "usability" scores were not as high, with 7.5 as the average score. The most common UI (user interface) criticism during our trip is that the data viewers were not translated to the local language, so we spent some time towards the end of the trip translating the data viewers and surveys to Bahasa Indonesia. Another common criticism is that the standalone data viewers had inadequate color keys. Audiences offered helpful suggestions on how to improve the color keys, especially since color is such a central component to OpenIR's usefulness.

Other important suggestions overlapped with some items on the OpenIR team's "to-do" list: implement better tutorials, perhaps via a how-to video; offer a smoother way for new users to build their own OpenIR deployments; and include data from a time range, which would be more useful for environmental monitoring.

From our observations and interactions, both during and prior to my Indonesia travels, we have seen many kinds of people respond to OpenIR. We think the project has most resonated with journalists, whether it be a professional or citizen journalist, an environmental or data journalist. We noticed this when two environmental journalists in Ethan Zuckerman's "News in the Age of Participatory Media" wrote about the project; when WNYC and Guardian UK data journalists became quite enthused about OpenIR at the Knight/Google/O'Reilly NewsFoo conference; and we noticed this repeatedly in Indonesia, as well. We think it's telling that our strongest community partner is an environmental journalist, and that some of our most meaningful meetings were with indigenous groups trained as citizen journalists. While expert geospatial professionals don't need OpenIR, and map neophytes don't easily understand OpenIR, the people trying to tell complex environmental stories find that OpenIR offers powerful images with which to underscore those stories.

Summative Surveys

General Summary

The Summative Survey was a 15-question survey hosted by SurveyMonkey.com. The survey was in both English and Bahasa Indonesia. From January to Marcy 2013, there were 61 respondents: 45 in English, 16 in Bahasa Indonesia.

While one of OpenIR's main purposes is to bridge the gap between infrared data and the non-specialist population, it seems that there is still a lot of work to be done. Most of the respondents commented on the lack of usability, difficulty to understand the data presented, and non-intuitive layout. It seems that a tutorial would help greatly, as well as a statement of purpose. The legend, while helpful, goes largely unnoticed by users for some reason. Again, the question arises of what people do with the information that OpenIR presents to them. It seems that 10 minutes is a cutoff point for people to lose interest/stop exploring out of frustration from not being able to understand the website.

Comments On Usefulness

- "No doubts about the usefulness of free IR data in risk management. Nevertheless, non expert people probably don't understand the meaning. A download option would be very important for experts, while clearer comments about what information the maps show would be desirable for non-expert users."

- "The different layers are useful, but it's not immediately clear what the data means — i.e. what does area of low moisture mean, does it suggest vulnerability?"

Comments On Usability

- "It's easy once you understand how it works."

- "Perhaps I am not the intended user. However, I don't understand how this map can be used or contributed to."

- "It seems that you pasted a lot of information onto a map. That's nice, but it would be great if you would give some tutorial how to actually use it."

- "The implementation of "layers" is confusing. Some behave as layers in Google Maps behave (for example, can display multiple layers at once) while in other cases, only the most recent layer is displayed. Also, the interface for drawing shapes is inconsistent (lines and shapes are created by clicking, circles by dragging) which was annoying."

- "In looking at the sites, not so clear to me immediately how to use them for beneficial results. However, I didn't study the applications and notes on the main site."

- "The interface has pretty low learnability (as a first-time user, I found it difficult to learn how to use the tool without instructions)"

Comment On Understandability

- "As I said before, open access to IR data is extremely important for expert users, but a common user couldn't understand the difference between different layers. Some sort of explanation is advisable."

- "As mentioned in usefulness, a map legend would be helpful. While the visualization is useful for general understanding, further explanation of the displays would increase the understandability."

- "I actually thought this was a bit of weakness. There is no interactive help to either a) explain how to toggle the layers, etc., or b) explain how the various layers were generated (doesn't need to be an algorithm level explanation, but some description of how the maps were produced from satellite imagery would give users additional insight/ trust). Users might also be puzzled by apparent linear boundaries in the mapped quantities, e.g. vegetation IR has a strong boundary that is nothing more than the border of a mosaicked scene – remote sensing users will understand that but non-technical users not so much."

- "It's good that there are not to many menus/items confusing the image. But at the same time, there's not enough information about the images available. Maybe a status bar on the left or right side that can be turned on/off would be useful?"

- "Might be nice to have a statement of purpose at the bottom or at the beginning."

Comments On Attractiveness

- "Very appealing look to the viewers."

- "The explanation box on the top-left could stand out more — didn't really notice it at first, even though it gives important info for understanding the software."

Comment on how OpenIR helped you learn

- "OPENIR is at this stage just too complex to actually learn anything from it. Even after 10 minutes of use I felt it was clumsy."

- "Unfortunately, I don't have the context to learn about this, just by poking around the website for 10 minutes."

- "Replete with useful information that is easy to find once you learn how to navigate the interface"

- "Comment on how you would use OpenIR with or to replace other software, and what other software did you use"

- "OpenIR would be useful when combined with ENVI for detailed analysis."

- "Used it in standalone mode. Can the layers actually be exported? I did not see any mention of that."

- "There are other platforms where ASTER data are implemented, which are more useful for me, such as NASA's World Wind Viewer, used by Geoscience Australia"

- "Do not currently use any software for which OpenIR would be a substitute"

Is there anything else you'd like to say about OpenIR?

- "OpenIR is a valuable tool for local populations and outside analysts/environmentalists. Continue the good work providing open access to IR and environmental data."

- "Looks like a cool tool – I'm not sure what it's for (as in, what would I do with the information I gather from looking at the maps) exactly, but maybe that's okay. Is it a technical tool or is anyone supposed to be able to use it? Also, I wasn't sure how to change the area I was looking at."

- "please please please work on usability. Having data online is nice, but you need to make it user friendly to actually make an impact."

- "if it had a more applicable use for me then id probably like it more"

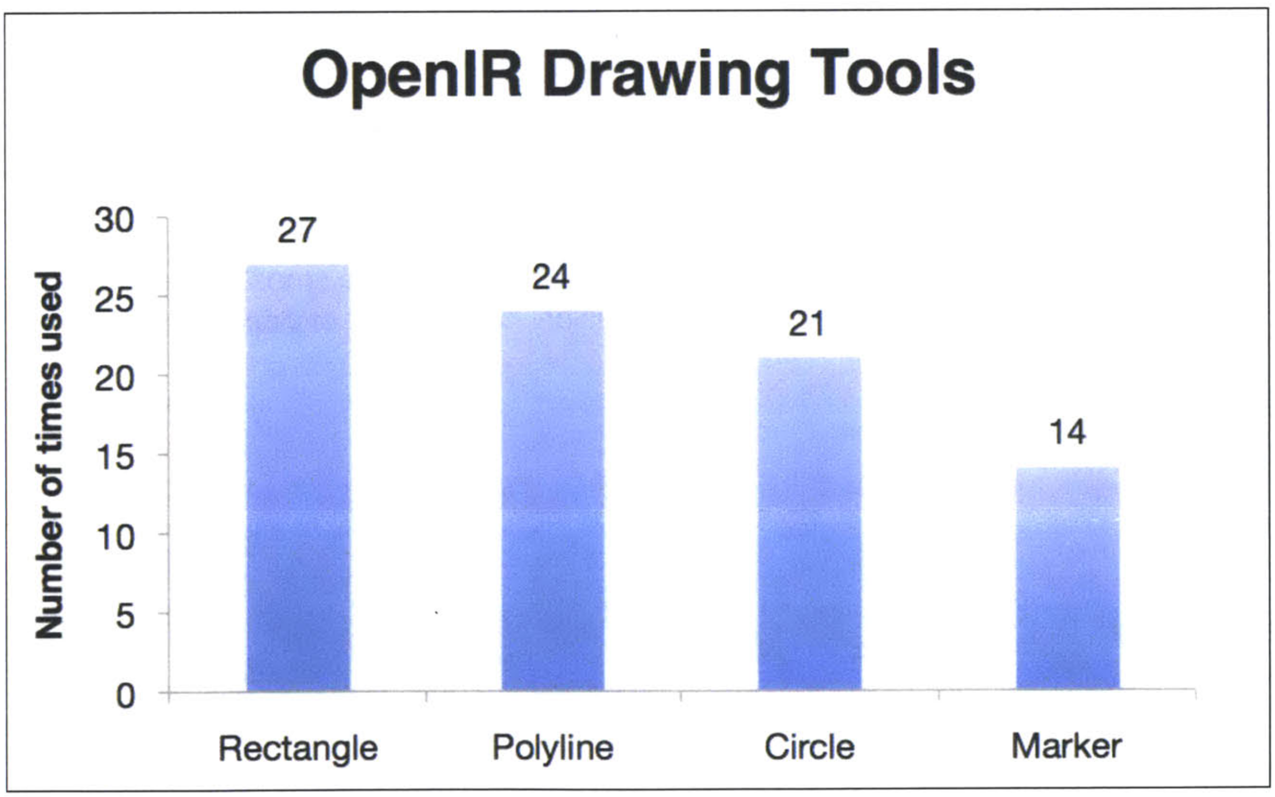
- "the maps load kind of slow, especially when zooming in and out."

Chapter 14: Summative Evaluation, Event Tracking
December 2012 to March 2013

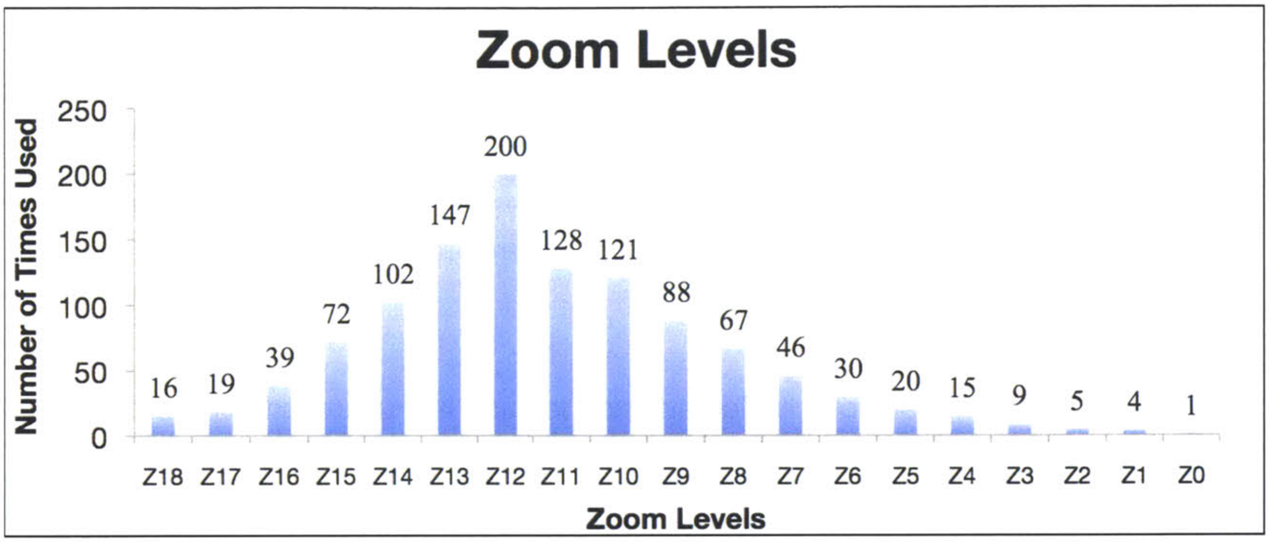
As OpenIR's Summative standalone Data Viewers were used over a four month period, all user-initiated events were tracked and written to a CSV (comma-separated value) file. These events included the use of drawing tools, zooming, activation and deactivation of IR layers. For all of these events, the IP address of the computer initiating each event was also tracked. There were slight differences between the tracking of the Jakarta and Java/ Kalimantan viewers, so we discuss them separately below.

Jakarta

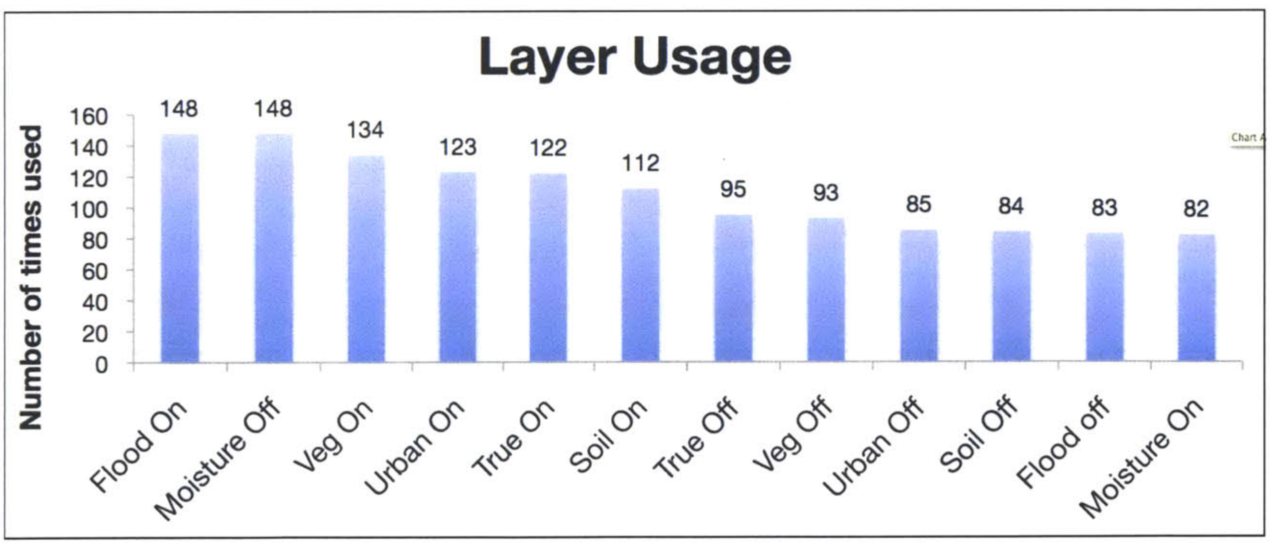
Of the drawing tools, the rectangle drawing tool was used the most. However, the circle and polyline tools were both used more than 20 times. There may need to be further thought put into the efficacy and purpose of the marker tool, as it is used significantly less than the other tools.

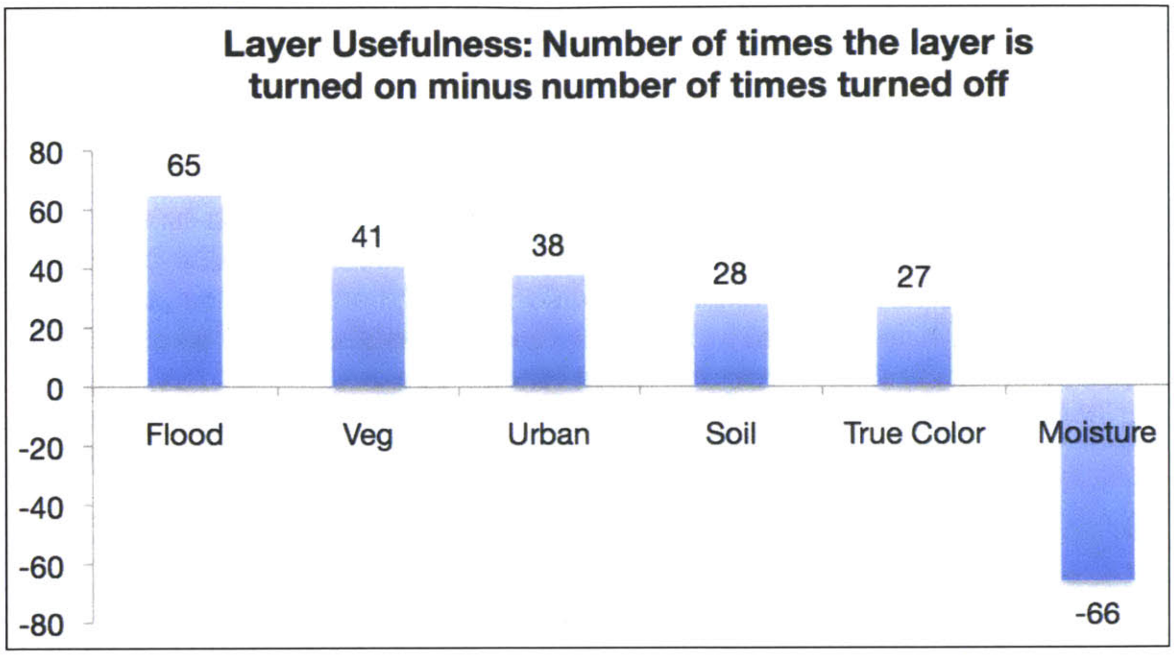


Of zoom levels, Zoom Level 12 is clearly used far more than any other zoom levels. However, this is likely to be merely an indication of where the default zoom level is. It is important to note that the further you get from the default zoom level, there is a dramatic drop in usage. Whatever we use the data viewers for, it may be important to have the default zoom level be the most pertinent zoom level.



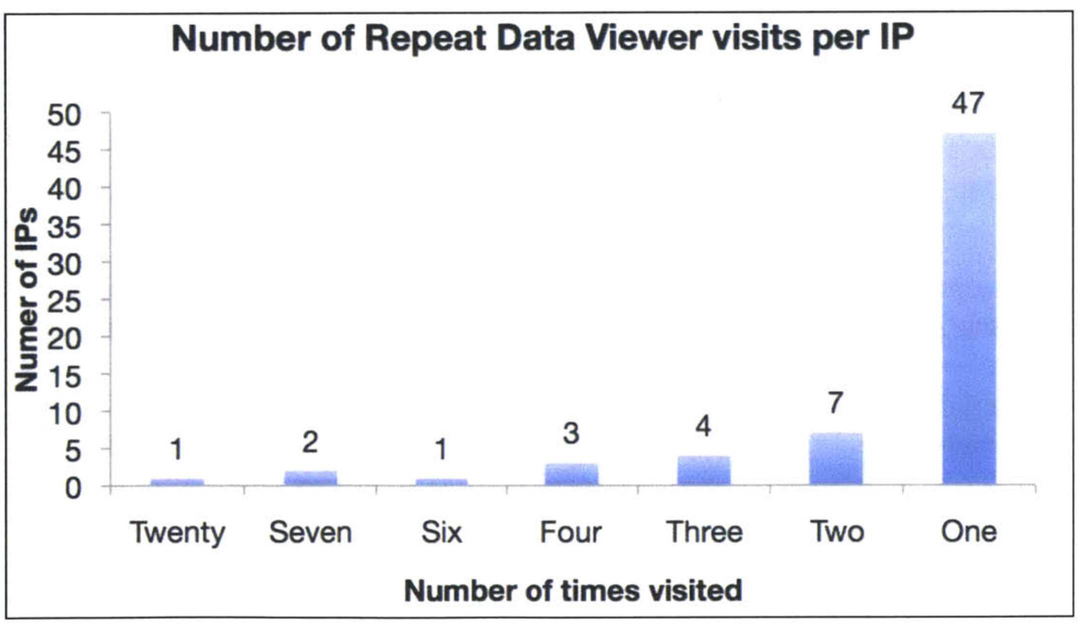
From the two graphs below, it seems clear that the flood layer is the layer of most interest to users. The moisture layer is the only one with negative "layer usefulness", most likely directly attributed to the fact that the moisture layer is the default layer displayed when the data viewer is loaded. Not counting moisture and true color, the soil layer is by far the least used layer.



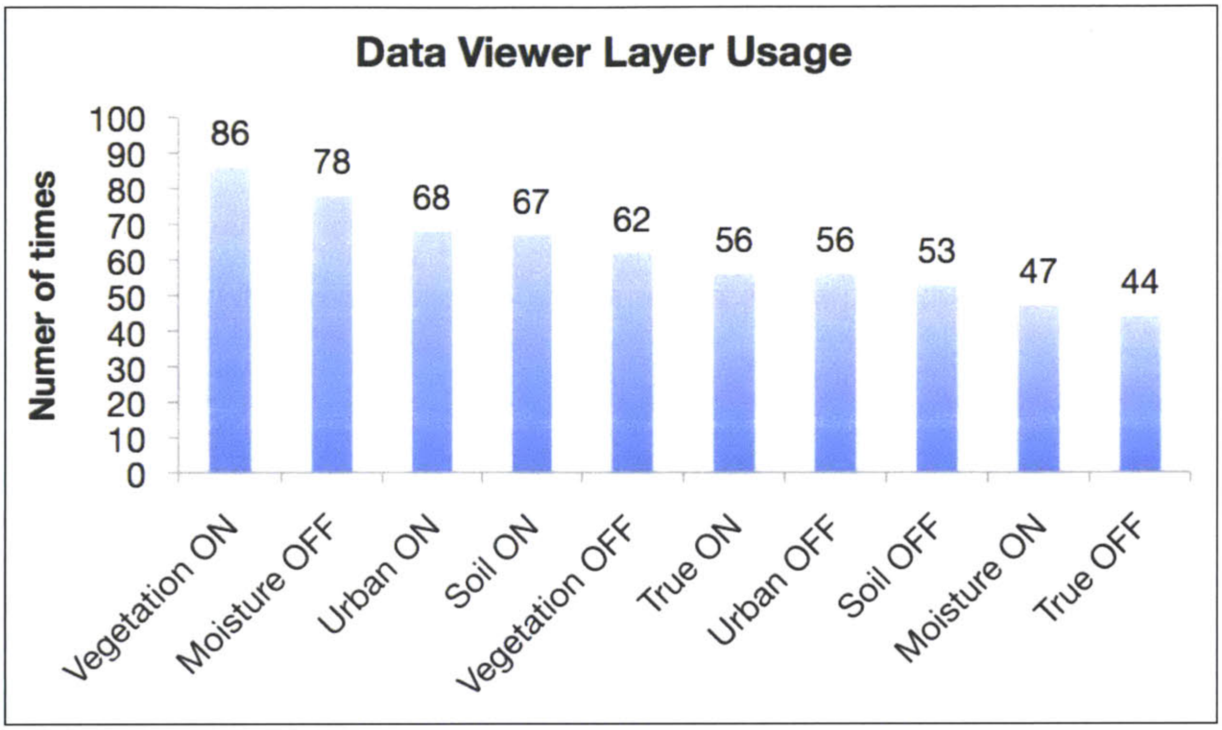


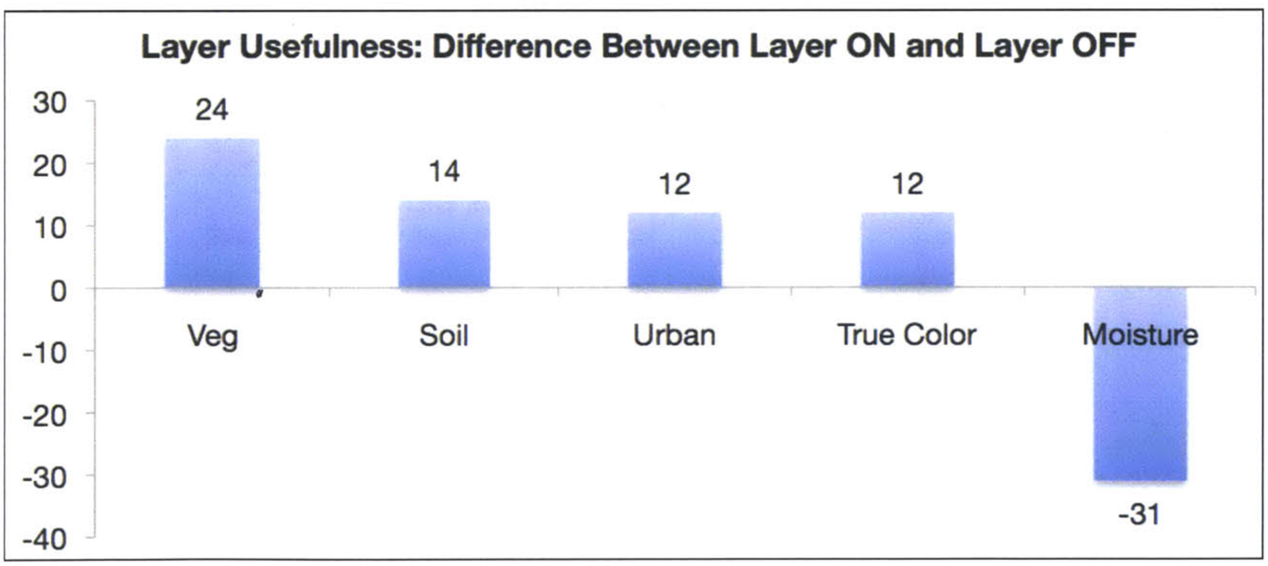
Java/Kalimantan

There was no data related to user usage available regarding drawing tools or zoom levels – only IP addresses, time used, and layer usage. In the layer usage, there is no flood layer usage data, as there is no flood layer data. It should be noted that IP addresses that were identified to be MIT IP addresses were taken out of this particular data analysis in an attempt to not include usage by OpenIR team members (i.e. during demonstrations of the software). However, this does remove a significant amount of data from the MIT populace not affiliated with OpenIR.



There is clearly an issue with retaining the interest of the user, with most users not returning after one visit.





Again, vegetation is the most popular layer (after flood in the Jakarta data viewer). The moisture level again is the only layer with negative "usefulness," but again that can probably be attributed to it being the default layer.

In conclusion, the usage concerns to address are retaining the interest of the user, determining a clearer purpose for the marker tool, and ensuring that the default zoom level be the most relevant zoom level.

These conclusions are based on a limited amount of data. It is also impossible to distinguish between what usage was by the OpenIR team and what usage was by external users, and the internal usage data may be skewing the results significantly.

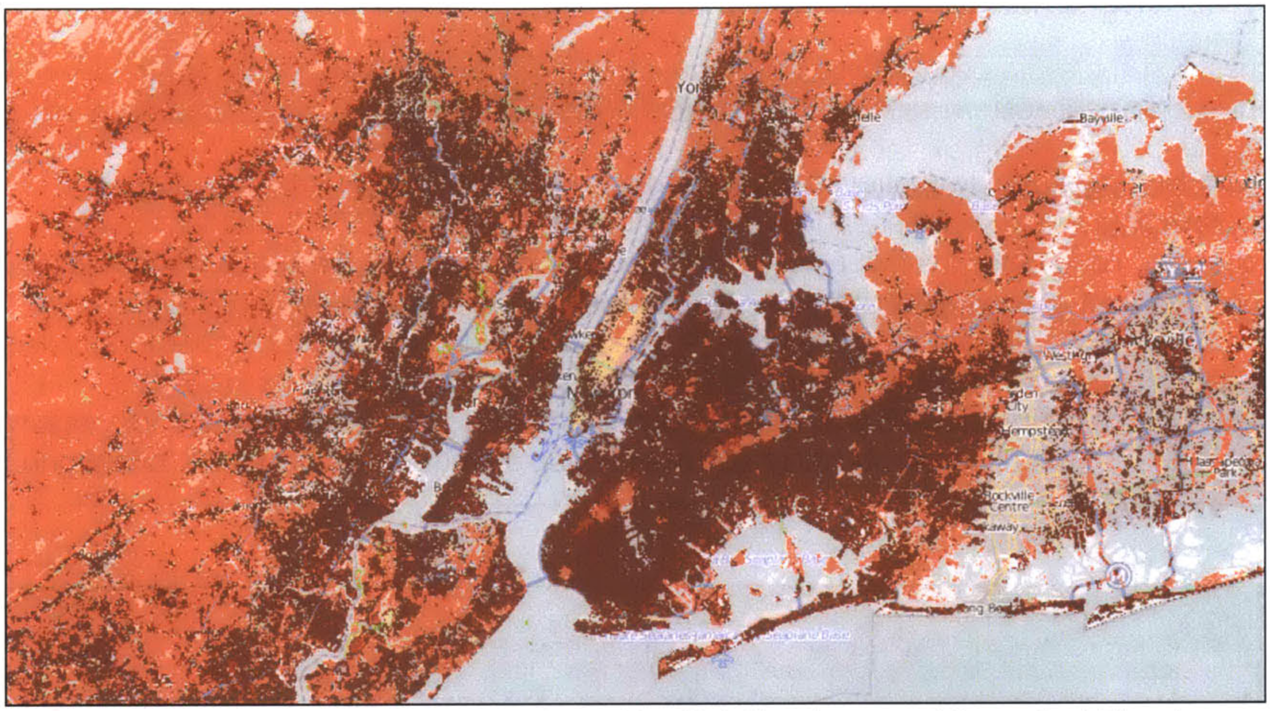
Chapter 15: Data Validation, Hurricane Sandy 2012 Flood

December 2012 to January 2013

After Hurricane Sandy destroyed many buildings in coastal New Jersey, New York, and Connecticut in late October and early November 2012, I asked FEMA's Ted Okada if observed flood data was available. FEMA Risk Analyst Jesse Rozelle sent a 3-meter (per pixel) dataset of flood extents for New York, New Jersey, Connecticut, and Rhode Island from early November. Undergraduate researcher Stephanie New, DuKode collaborator Ilias Koen, and I worked to prepare these datasets and merge them together in QGIS.

Hurricane Sandy, early November flood extents for New York, New Jersey, Connecticut, and Rhode Island. Data from FEMA.

Stephanie also used GDAL and ImageMagick to generate a Flood Risk Index Map for the New York area, based on the 4-index algorithm discussed in Chapter 6. In the Landsat imagery for this area, there is an obvious "no data" area on the map that cuts through Long Island. There are also some "no data" areas due to cloud cover.

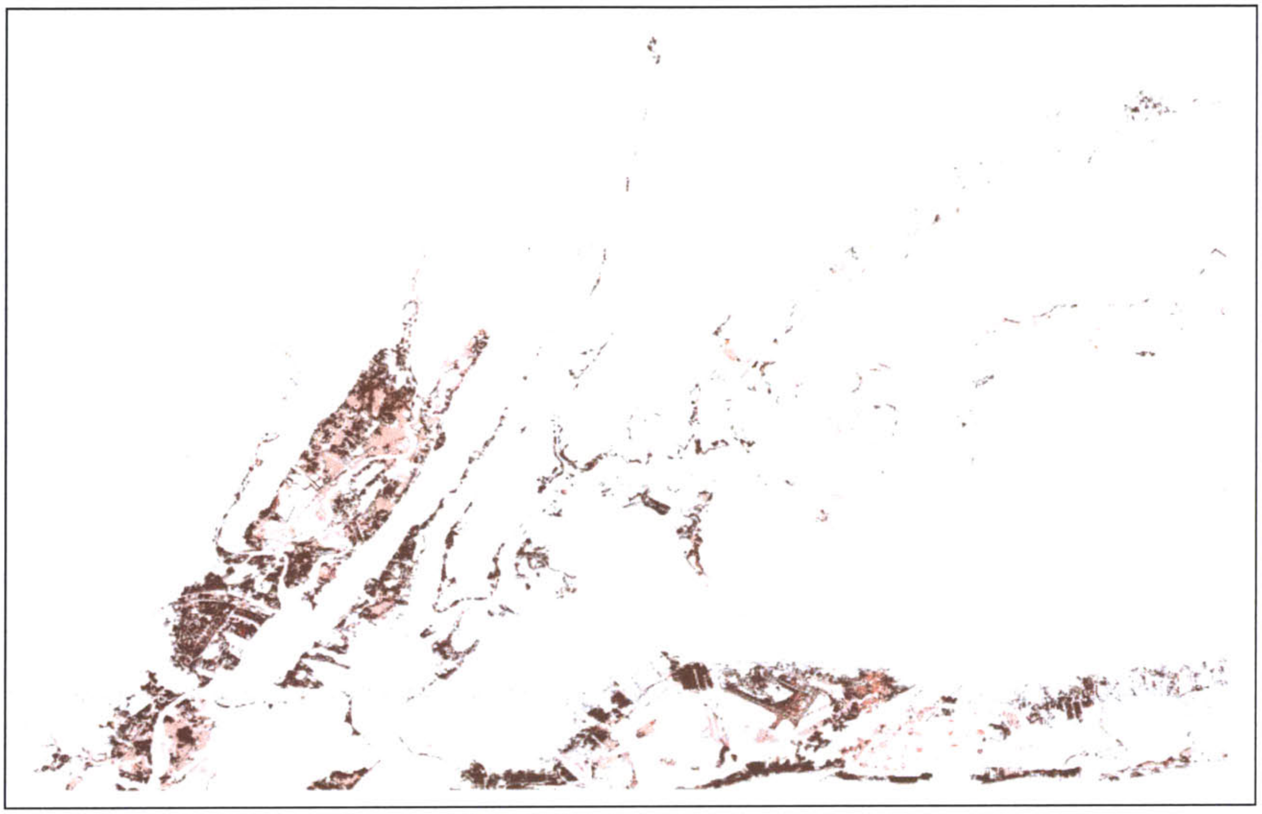


OpenIR Flood Risk Index Map for the "tri-state" area: New York, New Jersey, and Connecticut. There is an unresolved seam in the data (that looks like a zipper or railroad) near the right of this image.

Using QGIS, we output the intersection of the FEMA data with the OpenIR Flood Risk Index levels, using the same flood extent clipping mask process as discussed in Chapter 10. This process was conducted separately for each index level (urban low-lying, urban high-lying, vegetation low-lying, vegetation high-lying), so that area for each intersection could be calculated separately.

At left: color scheme for cross-section map below.

When each index level was processed, it was imported into a single QGIS project and colored with the same color scheme of the original Risk Index Map, with colors ranging from dark red to pink. The name of each layers reflects the processes conducted on it: The state flood data was unified (NY_NJ_CT_RI), then used to "clip" the index levels (urb_LL, urb_HL, veg_LL, veg_HL).



Intersection of FEMA flood extents and OpenIR Risk Index Map.

Similarly to the Jakarta 2007 validation, the results of the NYC validation indicate that urban low-lying areas, which the algorithm evaluates as areas of highest risk, have the most overlap with the observed flood extent data. For this Hurricane Sandy observed flood extents, 67.63916% of it overlaps with the OpenIR's urban low-lying layer (see table below).

However, the results of the NYC validation differ in that the next largest overlap between the OpenIR and observed maps is in low-lying areas of vegetation, which the OpenIR algorithm evaluates as areas of lowest risk. There is overlap of 26.17863% between the observed flood extents and the vegetation low-lying layer.

The difference between the Jakarta and New York flood validations may be indicators of different flood types.[82] While both areas are prone to coastal flooding and urban surface water runoff, Jakarta may have particular issues with ground failure from its clogged waterways.[83] New York may face more riverine (river-based) flooding, particularly flash floods.[84]

OpenIR New York Flood Risk Index Map, Areas Flooded in Oct/Nov 2012			
Index #	Feature	Color	% Overlap with 2013 Data
4	urban low-lying (urb_LL)	dark red	67.63916
3	urban high-lying (urb_HL)	red	1.81493
2	vegetation low-lying (veg_LL)	salmon	4.36728
1	vegetation high-lying (veg_HL)	light pink	26.17863

Beyond the scope of this thesis, we would like to re-run all of our validations, consult with flood risk experts, and then compare the validations against flood risk as determined by software including FEMA's Hazus and the World Bank's CAPRA.

[82] http://training.fema.gov/EMIWeb/edu/docs/fmc/Chapter%202%20-%20Types%20of%20Floods%20and%20Floodplains.pdf

[83] http://www.theatlanticcities.com/jobs-and-economy/2013/01/flooding-jakarta-city-swamped-its-own-success/4426/%5D

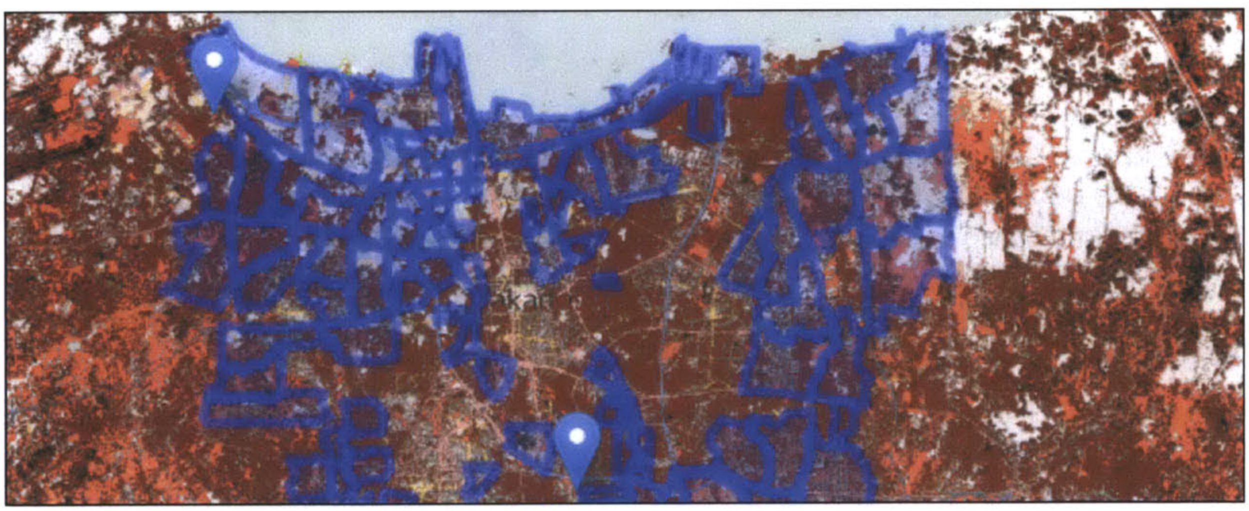
[84] http://www.nyc.gov/html/oem/html/hazards/weather_flashflood.shtml%5D

Chapter 16: Data Validation, Jakarta 2013 Flood
January 2013 to March 2013

My fieldwork in Indonesia was scheduled during MIT's Independent Activities Period in January 2013. This time of year is also the height of monsoon season in Jakarta, and heavy rain from January 16th to 18th flooded the city at levels unseen in decades, incapacitating a place used to major floods every five years.

During the floods, Kristy Van Putten of Australia-Indonesia Facility asked several local mappers to visualize flood current flood levels. She sent our team data of the flood from January 16th, 17th, and 18th. It was slightly different from the 2007 data we'd used: instead of showing flood extents, the data assigned a flood level to each neighborhood in the Jakarta metro area. Flood levels ranged from 0 to 300 centimeters for each day.

When I returned to MIT, undergraduate researcher Wendy Cheang and I conducted a validation process with this new data. We focused on January 18, the last day of the major floods.

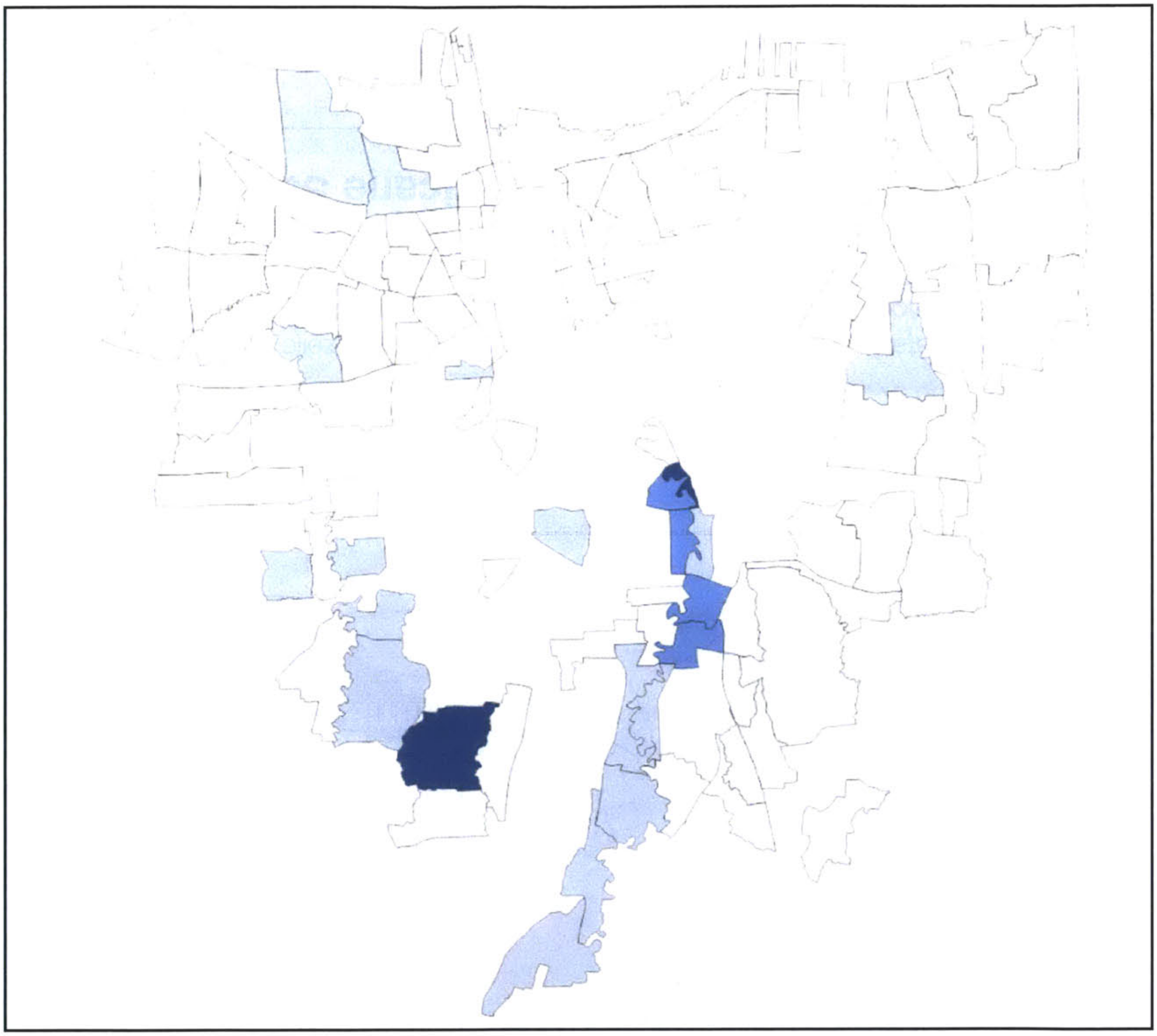


Jakarta's neighborhoods, outlined in blue, with the OpenIR Flood RIsk Index map as a base layer.

Validation Process

The validation process followed the same methodology as the Jakarta 2007 validation process, with one modification to accommodate the difference between flood extent data (Jakarta 2007) and flood level data (Jakarta 2013). Flood levels were divided into four classes: Class 1 (0-75 centimeters), Class 2 (75-150 centimeters), Class 3 (150-225 centimeters), and Class 4 (225-300 centimeters). Each OpenIR Risk Index level (Urban

Low-Lying, Urban High-Lying, Vegetation Low-Lying, Vegetation High-Lying) was then validated against these quartiles, making for a slightly more granular comparison than that of the 2007 data.

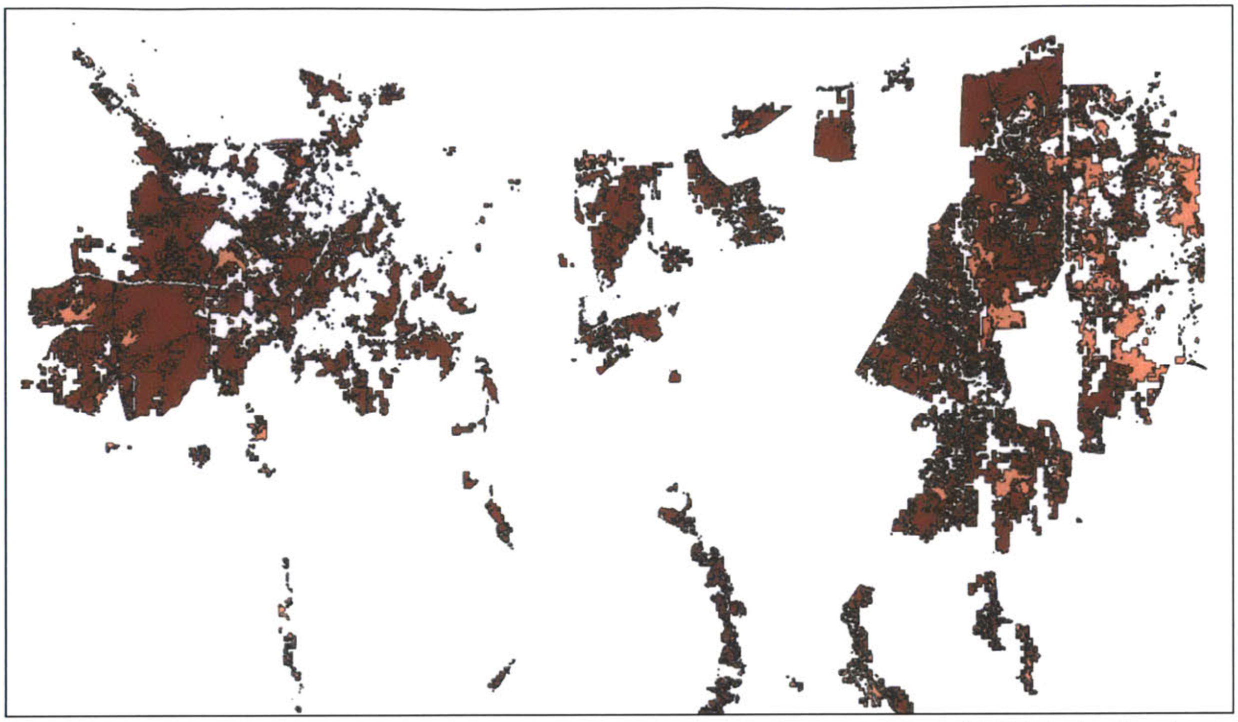


Jakarta neighborhoods, colored by classes flooding level for January 18. White is the least severe, while dark blue is the most severe.

Validation Results

The validation results show similar issues to the 2007 validation results: the majority of severe flooding (Classes 2-4) took place in urban low-lying areas, while the only other areas significantly flooded were vegetation low-lying areas. The vast majority of affected areas (93%) were in urban low-lying areas.

OpenIR Jakarta Risk Index Map, Areas Flooded in 2013		
	Percent (of total)	**Percent Subtotals**
Urban Low-Lying		
class1 (0-75 cm)	92.60688091	
class2 (75-150 cm)	0.70891378	
class3 (150-225 cm)	0.060280254	
class4 (225-300 cm)	0.027320897	
		93.40339584
Urban High-Lying		
class1 (0-75 cm)	5.26E-14	
class2 (75-150 cm)	0.005615677	
class3 (150-225 cm)	5.70E-06	
class4 (225-300 cm)	0.000387288	
		6.01E-03
Vegetation Low-Lying		
class1 (0-75 cm)	6.206812927	
class2 (75-150 cm)	0.361205955	
class3 (150-225 cm)	0.00985876	
class4 (225-300 cm)	0.003534004	
		6.581411646
Vegetation High-Lying		
class1 (0-75 cm)	0.006484228	
class2 (75-150 cm)	0.002454723	
class3 (150-225 cm)	0.000244903	
class4 (225-300 cm)	NONE	
		0.009183854



Jakarta neighborhoods affected by severe flooding, colored according to the OpenIR Risk Index four-level schema (i.e. urban low-lying is dark red, vegetation high-lying is pink).

These data validations highlight the difference between flood risk and the vulnerability of human in their built environment: vegetative low-lying areas are at greater risk for flooding, but depending on their proximity to urban areas, they may or may not affect human vulnerability. In modifying the Risk Index for future floods, it would be more accurate to simply show all low-lying areas as having higher flood risk than all high-lying areas; human vulnerability could be shown through a different, more complex algorithm that takes into consideration more inputs of the original directed graph algorithm (see Chapter 6).

Chapter 17: Implementation Summary & Future Work

Implementation Summary

To reiterate OpenIR's innovations:

- Programmatic interpretation of IR data through eco-feature highlighting and automated risk map creation.
- Tools and plug-ins for general audiences to annotate and validate satellite imagery with on-ground observations.

These goals were achieved, with an emphasis on front-end prototype usability; the back-end implementation is still at a rough stage. Looking back, the thesis implementation mainly took place as a series of software feature developments, in which small, "sandboxed" projects (Data Viewers or Ushahidi deployments) were built, each with different UI or map features, and sometimes with different geographic locales. UI feature developments included drawing tools, interactive keys, a streetmap overlay with opacity control. Back-end feature developments include IR tile acquisition and compositing, multi-tile image cleaning and compilation, automated risk index map generation, event tracking analytics, a Tile Map Service plugin for Ushahidi, and early stage spectrograph implementation.

The "sandbox" style of was useful, in that it allowed my collaborators and I to develop, experiment, and user-test many kinds of features without getting too deeply entrenched in any one type of feature or task. Having worked on OpenIR in this style for about a year, I now have a clearer idea of the literacy and software development issues with this project, and what kinds of features users need most.

Future Work

Beyond the scope of the thesis, it is important for this work to be more replicable; that is, a platform or installer package should exist so that any user can easily build an IR data viewer for any given region. My collaborators and I have developed a set of scripts to build these data viewers, but they need to be packaged for ease of use.

This would involve introducing atmospheric correction into our data processing algorithms, so that automatically processed IR imagery will have the most vivid colors possible. From there, it will be possible to develop an installer, with a user-friendly interface, that allows users to build an understandable Data Viewer for any region.

Because it offers many IR data sets as web maps, Google's Earth Engine project has played a role in OpenIR's implementation choices. But while Google Earth Engine plans to eventually release its IR data Javascript API, this release is not imminent. A few of OpenIR's stakeholders have raised the issue of processing global and/or national IR layers available via a Javascript library or the like. While OpenIR has focused more on city or regional data and the UI tools to make this data usable, I think it may be worthwhile to acquire global IR datasets from the UMD Global Landcover Facility, the current data source for OpenIR, and start processing them into WMS (world map system) or TMS (tile map service) tiles. However, there is no strong mandate for this effort, so my collaborators and I will pursue global processing based on available resources.

Future Collaborations

As the project expands, there are a number of stakeholders who have expressed strong interest in the further development of current Data Viewers for Jakarta, Java/Kalimantan, New York, and Makassar. These connections imply long-term collaborative relationships, and also imply future opportunities and directions for the OpenIR team and project.

AMAN & Harry Surjadi: OpenIR Installer

Harry Surjadi emerged as one of OpenIR's most committed community partners. He sees great potential in using IR data to tell environmental stories over time, and he has said that with his future work in disseminating his Information Broker model, he will use OpenIR as basis for visualizing collected information. In Harry's model, the Information Broker plays a more active role than the citizen journalist; he or she shepherds information from the citizen source to the decision maker. One of Harry Surjadi's original Information Broker systems is based at AMAN (Aliansi Masyarakat Adat Nusantara, or Alliance of Indigenous People of the Archipelago[85]) of Central Kalimantan.

AMAN's members don't just report information, they compile this information and use it to defend their rights in the decision-making process.

Because the natural environment plays such a primary role in the lives of indigenous groups, it makes sense that environmental imagery like that from OpenIR should play a role in the information they shepherd forth. Of particular interest to indigenous groups in this area are deforestation and forest degradation. When large scale deforestation occurs, as is often the case in corporate development, both indigenous livelihoods and forest biodiversity can be lost. The Dayak tribe is the predominant indigenous group in Kalimantan, and it has been greatly affected by deforestation issues. AMAN hopes to use indigenous boundary information gathered via GPS, in combination with the Information

[85] AMAN: http://www.aman.or.id/

Broker and OpenIR systems, to compile a robust set of visual documentation to support its efforts in land disputes.

During OpenIR's pilot work with Harry Surjadi, AMAN deputy chair Alfanius Rinting asked the OpenIR team how he could independently build and customize his own IR data viewers. Through the use of ethnographic interviews and GPS surveying, AMAN Kalteng is collecting current boundary and land use information from Dayak tribes across Central Kalimantan. This information will be an important tool for negotiations between the government, corporate plantations, and indigenous-rights groups like AMAN. But this tool will be strengthened if it can be coupled with Central Kalimantan environmental data, especially if that data has high spatial and temporal resolution. Independence is of primary concern to groups like the Dayak, which is reflected in both its land and data use.

After a conversation in which the team explained its current code, it became clear that a more automated, autonomous process was needed. The OpenIR team could take AMAN as a "client" and build data viewers for the group, but everyone agreed that this solution would be far from ideal in supporting indigenous independence. Currently, the most direct way of accessing IR satellite maps is through open government databases that require some geospatial expertise. Working with Harry Surjadi and AMAN, I would like to polish OpenIR development scripts into a turnkey IR Map Installer that offers groups like AMAN an intuitive method to build the data viewers it needs. Ultimately, this installer could offer all ecological interest groups a simple, powerful way to visualize the environmental issues that concern them.

Urban Risk Lab, NYS 2100, & FEMA: Flood Prep Mapkit

In the wake of October's Hurricane Sandy and January's severe Banjir (flood) in Jakarta, flooding is on the minds of people across the world, in both developed and developing regions. FEMA, having provided Hurricane Sandy flood extent data for OpenIR's NYC validation process, also talked to OpenIR about its new Open Government Initiatives. Working with this group and MIT's new Urban Risk Lab,[86] a "Flood Prep Mapkit" may develop as part of OpenIR. It would include

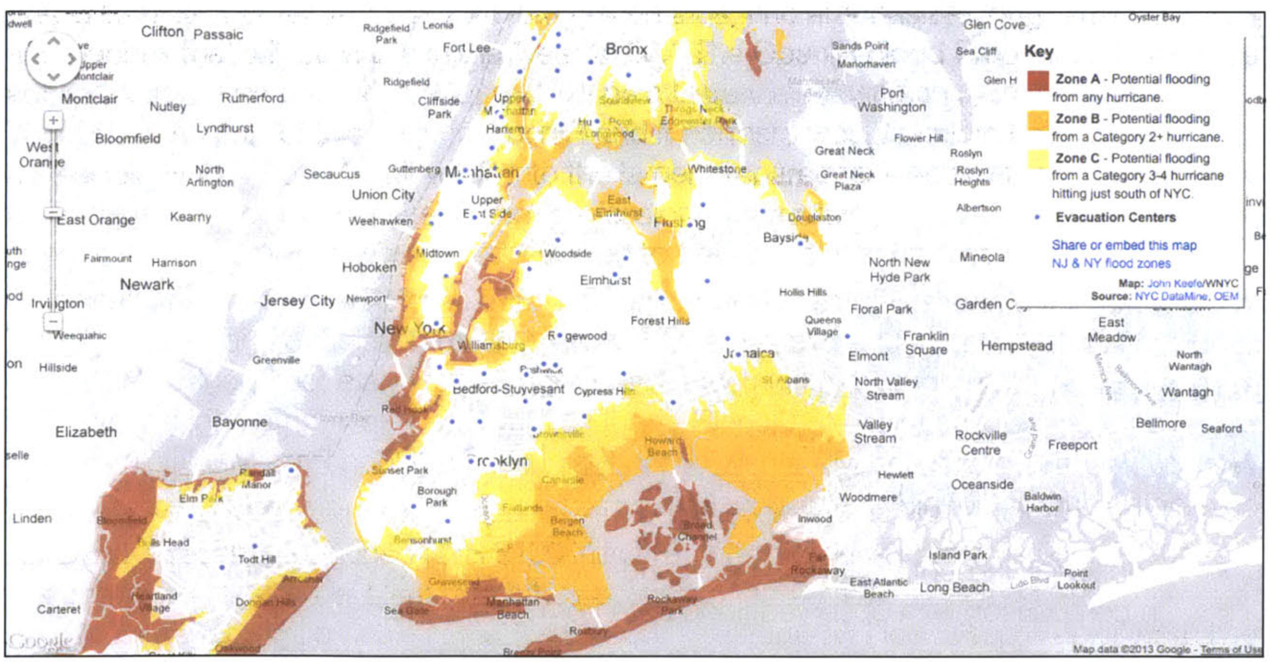
- additional, in-depth data validation and evaluation of Flood Risk Indexing processes in comparison with observed flood data from recent events
- porting of features from FEMA sources (HAZUS software, Flood Insurance Rate Maps, etc) to more open, web-map environments
- analysis of Flood Prep Mapkit with flood-focused social media and crowd map data

[86] MIT Urban Risk Lab: http://cau.mit.edu/overview/affiliated-labs

- dynamic evacuation route recommendations based on the Flood Risk Indexing process.

Professor Miho Mazereeuw, a founder of MIT Urban Risk Lab, is also a commissioner of New York's recently launched NYS 2100 initiative, which tackles problems of emergency preparedness in potential New York future scenarios, and emphasizes long-term approaches that will be viable in the "year 2100." The OpenIR Flood Prep Mapkit would be applied to the rubric of NYS 2100, and will be a Data Viewer with an initial basis of New York State. As part of this project, OpenIR and it partners would explore how simple, accessible Flood Prep Mapkit will fit into the larger preparedness ecosystem, from federal to local emergency management, from professional journalists to crowd mappers. As part of this project, protocols would be developed to better analyze the cross verification potentials of satellite+social data, and their derived products (i.e. risk maps) collected from OpenIR+Ushahidi deployments.

The Flood Prep Mapkit would supplement existing evacuation zone maps from public NYC Office of Emergency Management data. These maps show evacuation information for NYC only, and no public flood zone web map exists for areas surrounding the city.



WNYC Flood Evacuation map[87], based on NYC Office of Emergency Management data.

87 http://project.wnyc.org/news-maps/hurricane-zones/hurricane-zones.html

In the past year, it was highly useful to observe how bootstrap flood monitoring fit differently into the infrastructurally and economically developed New York City and the social media-dense, rapidly developing Jakarta. Each city was served differently by government, journalistic, and non-profit institutions. Going forward with the Flood Prep Map project, my collaborators and I would build a Jakarta Data Viewer with Flood Prep Mapkit and ask our extensive Indonesian community partner network to evaluate these maps for its applications in their emergency preparedness ecosystem.

Makassar Case Study and Time-Based Tools

Cycles of urban development and climate change reinforce each other, often in destructive ways. For example:
- increased greenhouse gas emissions from urban traffic are linked to
- increased outdoor temperatures, which in turn is linked to
- increased urban electricity consumption (i.e. for cooling).

Organizations like the United Nations Development Programme are trying to mitigate these destructive cycles by assisting smaller cities just entering into possible cycles of rapid development. One mode of assistance is its "climate vulnerability assessment," in which small cities can better understand their own environmental risks, and plan accordingly.

After returning from Indonesia, my collaborators and I started to work with John Taylor, founder of the Indonesia-based community mapping organization Solo Kota Kita, to develop an IR Data Viewer for the small Indonesian city of Makassar, subject of one of these pilot projects. Makassar is expanding its coastal area through "land reclamation," or the filling of coastal areas with soil from the nearby inland watershed. The OpenIR Makassar Data Viewer visualizes the state of land use to show how land reclamation may be tied to increased flood risk in this region.

The OpenIR Makassar Data Viewer offers the standard OpenIR infrared satellite map layers highlighting moisture, vegetation, soil type, and urban surfaces. John Taylor's team is combining OpenIR's data with an outline (shapefile) of the Makassar watershed, then extracting IR eco-feature colors (i.e. red=vegetation in OpenIR's "vegetation" layer) to show how lack of vegetation is increasing water run-off and flood risk. In John Taylor's words:

> "We want to show that the watershed is getting degraded, and how much so. If we can also show how fast, that would be very interesting. Using the OpenIR images we're going to be able to calculate how much of the watershed has vegetation/ permeable/ moist in 1991, then 1999, then today, and how fast it is getting degraded. This is crucial because it is the soil erosion that is creating flooding issues downstream."

Static IR maps have been useful visualization layers for this project, but the collaborative aims to deepen this work by introducing time sequences of Makassar's land use over the past two decades. To best show how Makassar has evolved over time, I would like to implement two innovative interactive tools that help users better understand this evolution:

SPECTRAL FEATURE IDENTIFICATION. For any given pixel in OpenIR's infrared imagery, surface spectrograms can be compared against entries in a spectral library (i.e. the ASTER spectral library), and a named feature can be identified, i.e. "construction asphalt," "dry vegetation," "sea water," etc. This ability allows for very fine-grained examination of environmental features. With this tool, in combination with an IR map time sequence, users will can visually see and textually read how land use of any area evolves over time.

ALGORITHMIC MOTION/COLOR AUGMENTATION: A team at MIT CSAIL (Computer Science and Artificial Intelligence Lab) has demonstrated a simple algorithm that uses augments color in a time sequence to reveal cyclical motion like blood flow and breathing.[88] The CSAIL team has made the algorithm open-source. My collaborators and I would apply this algorithm to time sequences of IR data to augment patterns in vegetation growth, soil movement, moisture change, and other cycles. Michael Rubenstein, a Ph.D. student on the CSAIL team, told me that this algorithm has never been applied to IR satellite data.

All of these tools will also explore alternative user interfaces for temporal data presentation. Like all of OpenIR's work, its Time Sequence tools will be open source, and all environmental interests, from UNDP to individual journalists and bloggers, can use these tools to track ecological change over time.

[88] Wu et al. "Eulerian video magnification for revealing subtle changes in the world." ACM Transactions on Graphics (TOG) - SIGGRAPH 2012 Conference Proceedings. Volume 31 Issue 4, July 2012 Article No. 65

Chapter 18: Evaluation/Validation Summary

Evaluations and/or data validations of OpenIR have been conducted almost continuously since the project started to significantly develop in June 2012. Every evaluation and validation process examined the project from a different perspective and shed new light on how to focus the project, yet no evaluation or data validation was satisfactorily conclusive. There are a number of likely reasons for this: internal evaluation by our team, as opposed to external evaluation by a third party; the broad engineering and user-centric scope of this project; and time limitations. Nevertheless, a multi-faceted approach to evaluating this project, through user responses and existing geographic observations, is an approach that we would encourage external evaluators to take.

Evaluation Summary

Front-End

The front-end interview questions were very broad, giving the OpenIR team a clearer picture of what general direction it should look to. Many of these interviewees were experts in the field of satellite imagery. However, interviewees that were both experts and interviewees that were unfamiliar with infrared satellite imagery largely agreed that the combination of satellite data and human data was extremely important. Potential usage suggestions included disaster scenarios, climate change, vegetative health, and human health. The front-end served as a confirmation that disaster scenarios and climate change should be part of the main focus of OpenIR in the immediate future.

Formative

The Formative interviews took place with individuals that were a part of our partner organizations in Indonesia (Air Putih, Humanitarian OpenStreetMaps, Knight International Journalism Program, University of Gadjah Mada, and RuaiTV), and FEMA. OpenIR had included by the time that these Formative interviews happened that OpenIR would bridge the gap between this data to non-experts, and so an easy-to-use user interface was crucial. We were able to get a great deal of feedback on how to make OpenIR easier to use for first time users.

The suggestions that were made and implemented in our Summative Viewer were:
1) legends/keys in the Data Viewer; 2) having instructions when hovering over something instead of having to click on markers in the data views; 3) a slider for pixels, allowing people to focus on specific risk levels ; 4) being able to see the layers when zooming in; 5)

a transparent map layer with street names to allow users to orient themselves; 7) using Indonesian Twitter keywords

Suggestions that were not added to the Summative Data Viewer were: 1) get a popular/influential person to be seen using the software; 2) ensuring the responsible use of social media when adding tweets to Ushahidi map; 3)switching up the green and red coloring on the vegetation layer to make it less confusing to view.

Summative

The Summative interview was originally made to be taken on a one-on-one basis with the interviewer in Indonesia. However, with the limiting factors of time and interviewing personnel, we did all of the interviews in group form. This meant that we skipped all biographical information gathering. This also meant that individual opinions that were held that may have been unpopular with the particular group being interviewed may not have been voiced, or if the person was less inclined to make their opinions heard, even if it was not a controversial opinion. There were also translation difficulties. In several of the meetings, we lacked a translator, and much information could have been lost in translation. After every meeting, the group members insisted that the same questions be emailed to them so that they could review the Data Viewers and questions at their leisure. However, no additional interview answers were provided via e-mail.

The Summative survey was distributed to the same population to try to counteract this. It was also distributed to populations that were not interviewed, such as people who have expressed interest in the development of OpenIR not in Indonesia, or people who were completely new to the OpenIR technology. Very few members of the original Indonesian study group responded to the Summative survey. However, surprisingly, most of the Summative survey respondents that were completely new to OpenIR were able to effectively navigate the Data Viewers and garner information on their own.

The possibility of winning the $50 Visa gift card could have played a major role in the English-speaking population's decision to take the Summative survey, but the surveys were sent to them only after that was publicized. However, the Indonesian surveys were mostly taken before the gift card was publicized.

Validation Summary

The strengths and weaknesses of the OpenIR Risk Index Map were highlighted in a validation against Jakarta 2007, Hurricane Sandy 2013, and Jakarta 2013 observed flood data. Insights include:

- There is a difference between flood risk and damage risk.

- Showing low-lying urban areas on a map can be a simple but effective way to convey areas where flood and damage risk intersect. It can show damage potentialities for all kinds of flood: river surges, flash floods, urban drainage, fluctuating lakes, and coastal flooding.[89] It may be least ambiguous and most helpful just to show this kind of feature as part of a risk map.

- Low-lying vegetation is a more complicated feature to evaluate: it is also at risk of flood, but it often (as in the case of mangroves and marshes) protects developed areas. More research is needed on how to best show this kind of feature as part of a risk map. In addition to vegetation, soil, natural irrigation land features may be included in a risk map, but we will need to implement research on how non-inhabited areas have best impact on developed areas.[90]

- Now that observed data has been combined with OpenIR maps, it may be useful to show this kind of historical observed data in the existing Data Viewers.

- Especially when they are processed rapidly, OpenIR Risk Index Maps may be a good companion for more complex, "offline" maps like FEMA Flood Insurance Rate Maps (FIRMs).

- Temporal IR data may be of greater use for deriving risk maps in future iterations of OpenIR.

---

89 FEMA, "Types of Floods an Floodplains." http://training.fema.gov/EMIWeb/edu/docs/fmc/Chapter%202%20-%20Types%20of%20Floods%20and%20Floodplains.pdf

90 Winsemius, H. C., et al. "A framework for global river flood risk assessments." Hydrol. Earth Syst. Sci. Discuss 9 (2012): 9611-9659.

Chapter 19: Conclusion

Considering the previous chapters' summaries of the implementation, evaluation, and validation for OpenIR, these are the main successes and limitations of the project to date.

Successes

- We demonstrated widespread interest in the combination of IR satellite maps with social and crowd data.

- We developed data viewers to display IR band combinations, sourced from NASA's Landsat program, for several different locations, including Jakarta and New York City, two highly populated regions with very different types of infrastructure. We developed a pipeline of GDAL-based scripts to output this kind of data viewer for any given locale. Even large areas like the island of Kalimantan can be addressed with this pipeline.

- We developed a Flood Risk Index algorithm based on risk assessments that use land classification and spatial combination techniques to input hazard/exposure and output possible impact. (Land classification is a common use of IR satellite maps.) These other assessments include Harris 2012, InaSAFE (AIFDR/BNBP), and Hazus-MH (FEMA). Since our algorithm uses IR-based land classification and elevation inputs, and since this kind of data is available for every landmass on Earth, this map can be replicated in data-challenged and developing regions.

- We built a plug-in to import IR and Flood Risk Index map layers into the participatory map system Ushahidi, and then deployed the system in New York and Jakarta to display crowd-sourced geo-tweets pertaining to floods (Hurricane Sandy and Jakarta Banjir). Both of these deployments were awarded as an Ushahidi Deployment of the Week.

- We built and tested additional tools to make IR maps more understandable. The work with spectrogram functionalities and spectral libraries are particularly potentially powerful.

- We kept the end-user strongly in mind, and conducted feedback sessions, collaborative workshops, surveys, and back-end analytics throughout the entire implementation process. To help with this, we built a large network of community partners in the U.S., Indonesia, and beyond. Some of the people in this network have become important allies in moving OpenIR forward, beyond the scope of the thesis.

Limitations

- Much of our data processing work is made redundant by Google Earth Engine, which is still a closed system that cannot be combined with other inputs (i.e. crowd data) via public server applications. This closed system goes against OpenIR's focus on accessibility. While large-scale data processing is a useful exercise, it is not the core competency of our team, and this redundancy may have kept us from progressing to more innovative ways of increasing IR data accessibility and literacy. It often felt like the innovations and insights were beyond the scope of this thesis.

- Though our data viewer pipeline allows us to rapidly compile an IR data viewer for any locale, it is still an unwieldy set of scripts that is best used internally. It is still not possible for most laypeople to output a data viewer based on our scripts, thus accessibility to IR data remains limited.

- The OpenIR Flood Risk Indexing algorithm is replicable for any locale, but its simplicity and limited inputs (elevation, moisture, vegetation, urban/impervious surfaces) make it useful only for preliminary and general applications, not for actual emergency management.

- Our validation process, which uses observed flood extents and flood levels, hasn't been validated in and of itself. Since our Risk Indexing algorithm calculates potential vulnerability rather than extent of damage, there may be more effective ways to validate our algorithm.

- The MIT Media Lab is a place where prototypes are made: projects that make for very useful demonstrations and teaching tools, while being too early stage to get much usage by the general public. Our standalone data viewers were not an exception. From our back-end tracking of usage and unique IP addresses, we saw that few people used the standalone data viewers more than once. On the other hand, we did not build a tracking system for the more complex Ushahidi deployments, which likely received much more usage than the standalone data viewers.

Next Steps

The unifying theme in the Chapter 17's three "Future Work" projects is the desire for an OpenIR software platform that makes major aspects of IR satellite data accessible and customizable for the layperson. This could be a web-based platform built on top of Google Earth Engine, or it could be a desktop-based installer that accesses, downloads, and compiles IR satellite maps into an upload-able data viewer. As of this writing, the latter looks more likely.

All of the community partners from the "Future Work" projects have expressed needs for time-based environmental imagery, especially for slower-onset crises like deforestation and watershed degradation. In future OpenIR applications, imagery from previous years and decades will be made available.

Finally, OpenIR has focused on IR data processing and the building of simple tools and algorithms for understanding IR satellite data. Particularly for the combination of social and satellite data, it makes a lot of sense to turn over the collected data to scholars who can analyze and search for patterns with greater skill. Our colleagues at MIT may help to this end, as may our new colleagues at Columbia's Tow Center for Digital Journalism and NYU's Center for Urban Science and Progress. As OpenIR transitions from a thesis into a longer-term project, we look forward to seeing how it may impact communities around the world.

Postscript: Project, Place, Process

Project Reflections

On the last day of Cesar Hidalgo's "Data Centric Projects" class at the MIT Media Lab in December 2012, students presented a semester's worth of work with datasets relevant to their personal research. I presented work on validating OpenIR's Risk Index Flood Map against Jakarta 2007 observed flood data (see Chapter 10). Ethan Zuckerman, one of the class instructors, responded to this presentation by pointing out that I have become very skilled at promoting OpenIR's positive aspects, but this data validation work marks the importance of showing the difficulties in the OpenIR project.

It can be quite a challenge for a researcher to take a balanced approach to his/her project, particularly at a hybrid place like the MIT Media Lab, in which research projects are all quite different from one another, and work that is entrepreneurially promoted tends to survive in the press-saturated environs of the lab. But a project that disregards its vulnerabilities and tensions will neither be robust nor useful. There are several axes of tension for OpenIR, at least as a Master's Thesis:

<u>Engineering-Centered versus Human-Centered</u>: As evidenced by OpenIR's validation/evaluation process, OpenIR was both an engineering and a human-centric project, as loosely defined by MASCOM (Media Arts and Science Committee, the academic faculty group of the Media Lab). Even though I originally thought of OpenIR as a human-centric project, it was placed in the "engineering" room on the Media Lab's Crit(ique) Day, and many of OpenIR advisers suggested that some of OpenIR's work needed to be validated against existing engineering processes. I think these suggestions, and the process of data validation, helped to strengthen OpenIR. Nevertheless, taking both engineering (data validation) and human-centric (usability studies) approach may have been too ambitious to be robustly represented by OpenIR.

<u>Research versus Advocacy</u>: Not only is there a tension between entrepreneurship and research, there is a tension between social advocacy and research. It can be difficult for a researcher to maintain neutrality while he/she uses his work to advocate for change amidst a disadvantaged group. Limited funding, time, and resources for these kinds of projects can also impact the neutrality of the researcher. With OpenIR so far, I think I've been a stronger advocate, both for the project and the populations it serves, than researcher. Underlying this bias is a deep commitment to the basic tenets of the project (i.e. accessibility of environmental mapping) and an interest in seeing this project survive beyond the scope of this thesis. However, I managed the project and so that most of my collaborators were more specifically focused on implementation, evaluation, and validation, and I think this, as well as thesis requirements for all of these stages, helped strengthen OpenIR as a research project.

Back-end versus Front-end: Data processing of IR tiles, sometimes in multiples and always at multiple (web map) zoom levels, was a useful, but perhaps needless task, especially considering the existence of Google Earth Engine, a project processes huge amounts of IR satellite data, albeit in a small, private group of beta testers. Because Earth Engine's API is not public, my collaborators and I could not work with the data in a publicly accessible way, necessitating that we develop our own processes for processing the data. Because it can be dangerous to rely too much on an external project, my collaborators, advisers, and I agreed that in-house data processing was quite useful. Nevertheless, spending so much time on data processing issues slowed the development of useful user interfaces, especially for our trip to Indonesia.

Short-term versus Long-term crisis: From 2005 to 2011, I developed earth and biogeographical visualizations with IR satellite data for years as part of the Science Bulletins group at the American Museum of Natural History. Most of these visualizations use IR data to make the same basic point: that climate change is rapidly accelerating and human behavior must be modified to this end. However these visualizations did not include data collected at the human scale, and this was part of my motivation with OpenIR—to better intersect IR environmental data with human data. Environmental crisis, particularly with the rise of participatory and crowd maps, seemed like the most relevant application of this intersection. However, crisis requires just-in-time data, and since the environmental scale changes more slowly than human data, most public IR environmental data is not collected more than once a day. There is a tension between short- and long-term crises, particularly because of the tension between the human- and the environmental-scales. This is not necessarily a problematic tension that OpenIR must solve, but it is an important one to acknowledge.

Personal Reflections

After the OpenIR travel, I spent ten days in the Philippines visiting family, reflecting on my time in Indonesia, and considering my identity as an Asian-American. While I've traveled throughout the North America, South America, Europe, and Asia for conferences and workshops, this was a unique experience, exposing my work and self to colleagues and family in much deeper ways. I am used to translation issues and cultural sensitivities. But as a 32-year-old woman visiting societies that are often more traditional and conservative than that of the Northeast U.S., I think I am entering a period in which being unconventional, particularly unmarried and childless, can be a cause for discussion. Harry Surjadi told me that he once interviewed Jane Goodall, and that Diane Fossey spend much of her research time in Kalimantan, eventually marrying a Dayak. I wonder how these women, as they progressed through their lives and careers, navigated the lines of culture, age, and expectation between very different societies. Fortunately, my "New York best friend" Crista Grauer, almost 75 years old and never married, has shared much of her wisdom and humor with me on this topic. She helped remind me that there are conservative societies in the U.S. too, and that these topics are never an issue with people who are true friends, no matter where you are.

OpenIR marks the first time I've worked extensively outside of the U.S. I think I had reasonable expectations as to the impact that our team and work would make: with any project, it never makes sense to prescribe a solution before better understanding the situation and the current solutions used. Add travel and cultural shift to this process, and in the end, I really did feel that this trip would best be approached as a series of introductions to communities and sub-societies. This trip is the first step of a long-term research effort, and it was a wonderful bonus that we found some community partners who are enthusiastic about collaborating with OpenIR over time.

Even though I am very conscious about being a person of color, my teammate Barry pointed out something that I think Western-trained people tend to forget: everyone, everything, every project, every nation, lies on a path to progress. It's naive to think that problems are solved when the "developed" or the "majority" help the "developing" or the "minority." Even the most developed groups need improvement. Even the most underdeveloped groups may be employing elegant, savvy solutions. This was very much the case throughout our travels in Indonesia, as we witnessed all of our community partners making ingenious use of all kinds of information/communication technologies. Our partners had a lot to teach us, and we were humbled to share our project in return.

Concluding Thoughts

In the past few years, I've come to realize that I tend to be attracted to projects focused on making large, esoteric, public datasets (i.e. genomic data, infrared satellite data) more accessible and understandable. I tend to approach these projects through a limited institutional approach, either by starting my own organization (i.e. The DuKode Studio) or by working on the project as a student. However, I've learned through the course of a startup, a design studio, and studies for a second Master's degree that, usually, big data projects require institutional resources. I'm grateful to have helped establish OpenIR in some small part at MIT, and I hope to find a way for the project to have some kind of stable institutional foundation.

I don't know if OpenIR will become a game-changer in environmental mapping and monitoring; there are so many axes to this effort, and I'm still trying to understand all of them. But I will do my best to continue OpenIR as a project, and hopefully, soon a platform, in that difficult but necessary space between private and public domains.

Glossary: Terms and Acronyms

API	Application Programming Interface
AIFDR	Australia Indonesia Facility for Disaster Reduction
AMAN	Aliansi Masyarakat Adat Nusantara: Indigenous People's Alliance of the (Indonesian) Archipelago
AMNH	American Museum of Natural History
Apache	An HTTP web server software program
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer, a high resolution imaging instrument that is flying on the Terra satellite
Band	In terms of satellite imagery, a region of the electromagnetic spectrum
COUHES	MIT's Committee on The Use of Humans as Evaluations Subjects
CSV	Comma-Separated Value (file format)
ESRI	Geographic software company (originally Environmental Systems Research Institute)
FEMA	Federal Emergency Management Agency (USA)
FTP	File Transfer Protocol
GDAL	Geospatial Data Abstraction Library
GeoTIFF	TIFF image file with geo-referenced metadata
GitHub	A service for storing files tracked with the Git source code management system
GLCF	University of Maryland's Global Land Cover Facility
GPS	Global Positioning System
GUI	Graphical User Interface
GWOB	Geeks Without Bounds

HOT	Humanitarian OpenStreetMap Team
ICT	Information Communication Technology
IKONOS	a commercial earth observation satellite, and was the first to collect publicly available high-resolution imagery at 1- and 4-meter resolution
ImageMagick	A command line image processing tool
InaSAFE	Indonesia Scenario Assessment for Emergencies, a plug-in for the software QGIS
iOS	iPhone/iPad/iPod Touch operating system
Landsat	Land Satellite: the longest running enterprise for acquisition of satellite imagery of Earth, administered by NASA
Leaflet	A lightweight Javascript library for displaying interactive maps on the web
MODIS	Moderate-resolution Imaging Spectroradiometer, a scientific instrument launched into Earth orbit by NASA in 1999 on board the Terra satellite
MySQL	An open source database system
NASA	National Aeronautic and Space Administration (USA)
OpenLayers	An extensive Javascript library for displaying interactive maps on the web
OSM	OpenStreetMap, an open source database of interactive street maps for the web
QGIS	Quantum GIS, an open source geographic information system software with a graphical user interface.
REDD+	Reducing Emissions from Deforestation and Forest Degradation. REDD+ is the next generation of REDD protocols.
SBTF	Standby Task Force, an international group of crisis map deployers
SDK	Software Development Kit
SMS	Short Messaging Service, a text messaging service component of phone, web, and mobile communication systems

Tile	An image systematically collected by a satellite
TMS	Tile Map Service, a web map image protocol
UNDP	United Nations Development Programme
UNOSAT	United Nations Institute for Training and Research (UNITAR) Operational Satellite Applications Programme
URL	Uniform Resource Locator
USGS	United States Geological Survey
Ushahidi	A popular crowd map software and service
WMS	Web Map Service, a web map image protocol

Appendix A: Awards and Press

OpenIR and its team members were recipients of the following awards:

- Alva Foundation Fellowship, 2013
- Ushahidi Deployment of the Week: Jakarta Banjir, 2013
- MIT IDEAS Global Challenge, 2012
- Geeks Without Bounds ATT Hacker Prize, 2012 (for unaffiliated hackers)
- MIT PSC Fellowship, 2012
- MIT Rodwin Award, 2012
- South By Southwest Dewey Winburne Public Service Award, 2012
- Knight News Challenge Finalist, 2012
- Ushahidi Deployment of the Week: Hurricane Sandy, 2012

OpenIR was profiled in the following news outlets:

- Engineering for Change (https://www.engineeringforchange.org/news/2012/09/24/how_to_manage_a_crisis_with_open_source_infrared.html)

- WiredUK (http://www.wired.co.uk/magazine/archive/2012/11/start/satellite-maps-but-better)

- BBC Producer (http://partnews.brownbag.me/2012/05/22/old-satellites-new-tricks/)

- Thomson Reuters Writer (http://partnews.brownbag.me/2012/05/22/a-trip-to-the-pacific-mapping-invisible-environmental-risks/)

- "The Buzz" (http://www.youtube.com/watch?v=zFf6MNMJYEA)

- SXSW Interactive Scholarship/Dewey (http://sxsw.com/interactive/awards/scholarship, http://sxsw.com/interactive/awards/dewey-awards)

- Ushahidi Deployment of the Week (http://blog.ushahidi.com/2013/01/24/weekly-openir-in-jakarta-windows-8-app-uchaguzi/)

Appendix B: Jakarta 2007 Validation Steps

The Jakarta 2007 Validation process consists of the following steps, which were then generalized and replicated for the NYC 2012 and Jakarta 2013 validation processes.

1. Download InaSAFE, QGIS, and any required dependencies.

2. Add the InaSAFE layer in QGIS: Click Add Raster Layer on the tool bar, and select inasafe_data-master>hazard>Flood_Current_Depth_Jakarta_geographic.asc.

3. Polygonize the InaSAFE layer: On the top toolbar, select Raster>Conversions>Polygonize (Raster to Vector), and input the InaSAFE flood raster layer.

4. Load one of OpenIR's Flood Risk Index levels (i.e urban low-lying): Click Add Raster Layer and select a layer of OpenIR's Flood Risk Index (the naming convention is "urb_LL"). Temporarily deselect the InaSAFE Flood Vector Layer.

5. Remove the background fill of the raster layer: Right click on the layer and select Properties. In the Transparency tab, click the Add Values from display button and click on the black background of the image. Delete the preexisting transparency row in the table (Indexed Value: -32768 Percent Transparency 100.00).

6. Clip the raster layer using a mask: Navigate to the top toolbar, and select Raster>Extraction>Clipper. Under Clipping mode, choose mask layer and select the InaSAFE Flood Vector Layer.

7. Repeat Steps 2-5 for other index levels (urb_HL, veg_LL, veg_HL).

8. Remove the borders from the clipped images, so that the images can be correctly filled:

    a. Open the clipped shapefile.

    b. Transform it from Polygons to Lines: Vector > Geometry Tools > Polygons to Lines.

    c. Toggle Editing and minimize the outer border to a tiny speck. Select the Move Features button and move the entire border away from the actual data. Choose Select features by freehand to select the border. Click Cut Features.

    d. Convert this layer back to polygons: Vector > Geometry Tools > Lines to Polygons. The tiny speck should be removed.

9. Calculate area percentages: Navigate to the top toolbar, and select Vector>Geometry Tools>Export/Add Geometry Columns. Use the Select Features by Freehand to select the entire raster layer. Then, click Layer>Open Attribute Table. Copy the table (Ctrl+C) and paste it into a spreadsheet editor. The second column of the table is the areas of the polygons that make up the layer, so auto-sum the second columns to calculate the total area of the intersection between OpenIR's Risk Index layer and InaSAFE's flood data.