Saarland Informatics Campus SIC

UNIVERSITÄT DES SAARLANDES

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

From Satellites to Social Media
What Data Tells us About Society

10. June 2025

Ingmar Weber

What I do: Societal Computing

The computer-enabled *study of societal phenomena*, and the use of computer-enabled methods to *support social development*

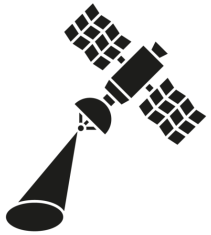
Computing *of* Society:

I use non-traditional data sources to measure und understand phenomena such as digital gender gaps, international migration, poverty, and more.

Computing *for* Society:

I work with stakeholders to support them in their missions with data insights, mostly in the global development and humanitarian context.

Talk Outline



- What do satellite images tell us about society?
  - Can you see tax evasion? Or internal displacement?

- How can we use advertising data for research?
  - Can we use it to measure migration? Or digital gender gaps?

- New opportunities come with old risks.
  - What are the dangers of “measuring society”?

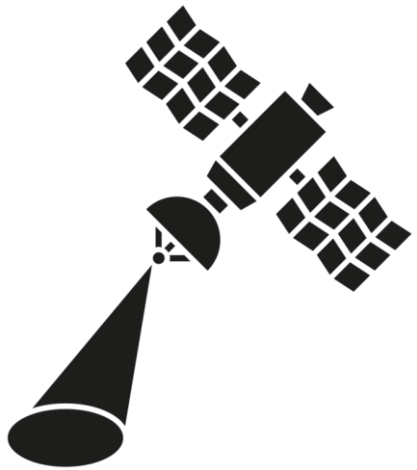
Fantastic Collaborators

Theophilus Aidoo (Saarland U.), Masoomali Fatehkia (QCRI), Manuel Garcia Herranz (UNICEF), Kiran Garimella (Rutgers), Karri Haranko (Aalto University), Ridhi Kashyap (U. of Oxford), Till Koebe (Saarland U.), Douglas Leasure (U. of Oxford), Joao Palotti (Scale AI), Francesco Rampazzo (U. of Oxford), Vedran Sekara (UNICEF), Emilio Zagheni (MPI for Demographic Research), …

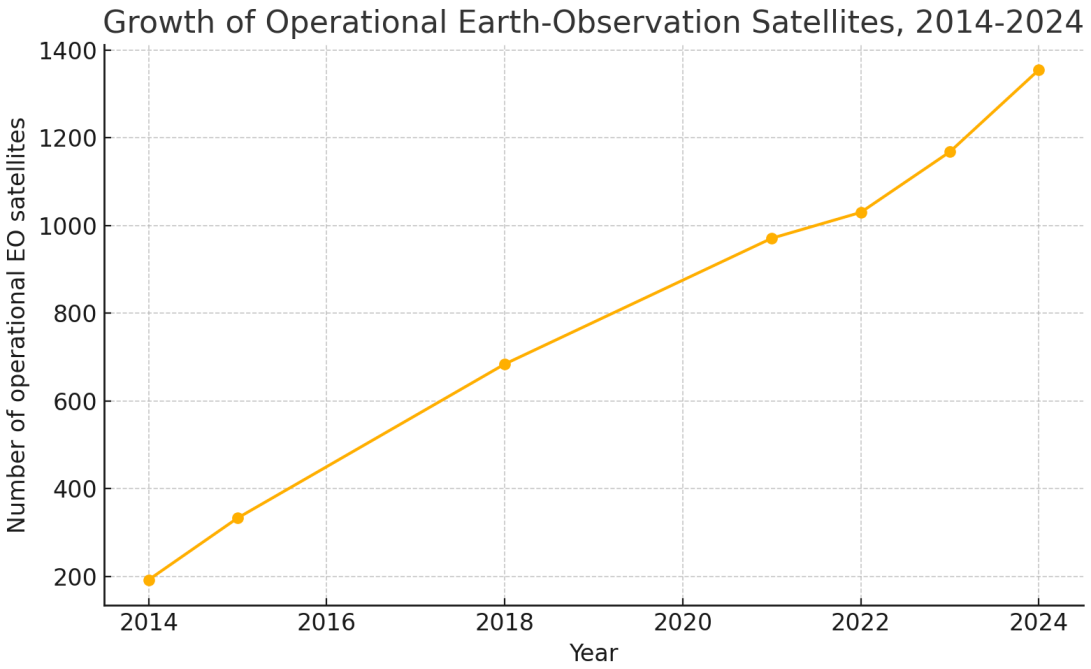
+ more collaborators at Saarland University, QCRI, UN agencies, NGOs, …

See list of key references at the end. Slides will be posted online.

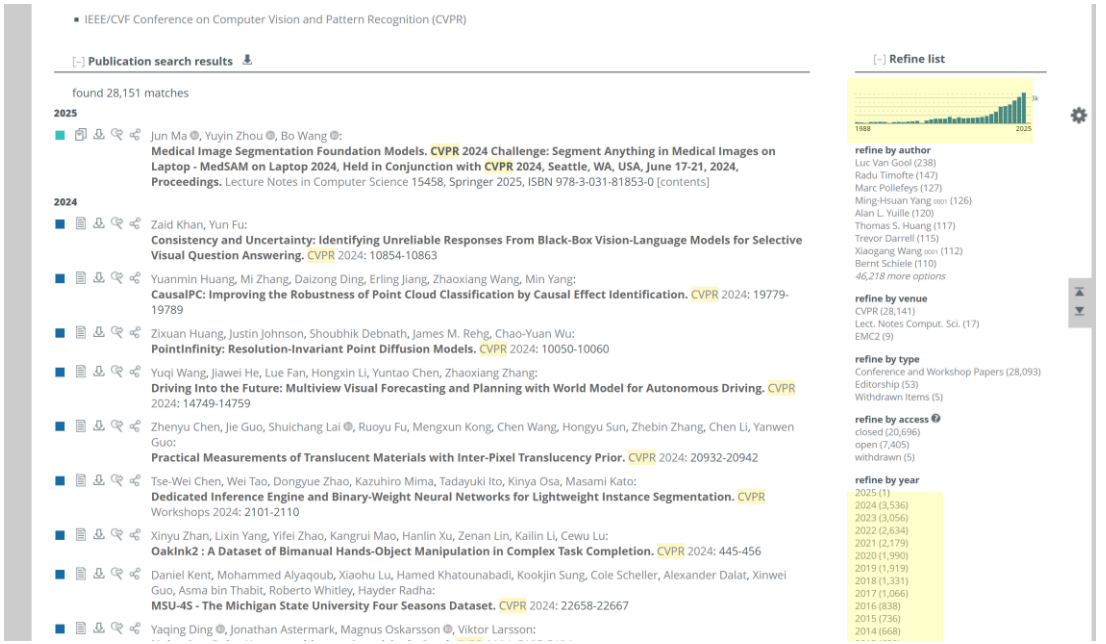
Satellite Images



Rapidly Increasing Capabilities



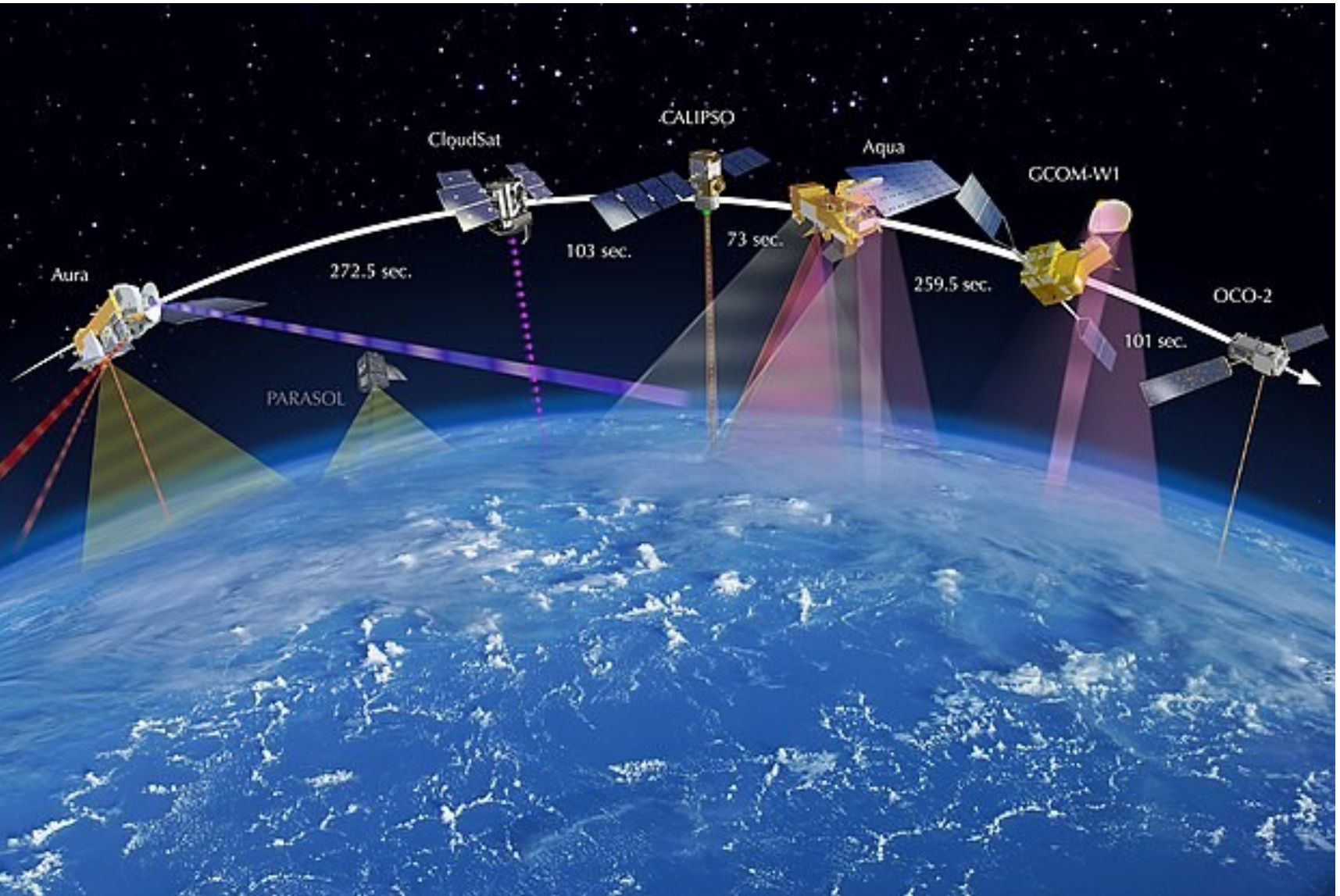
+



https://www.pixalytics.com/earth-observation-satellites-2023/
https://www.ucs.org/resources/satellite-database

https://dblp.org/search?q=venue%3Acvpr

Earth Observation Satellites (EO)



> 1000 EO satellites

30cm - 20m resolution

High-res is expensive

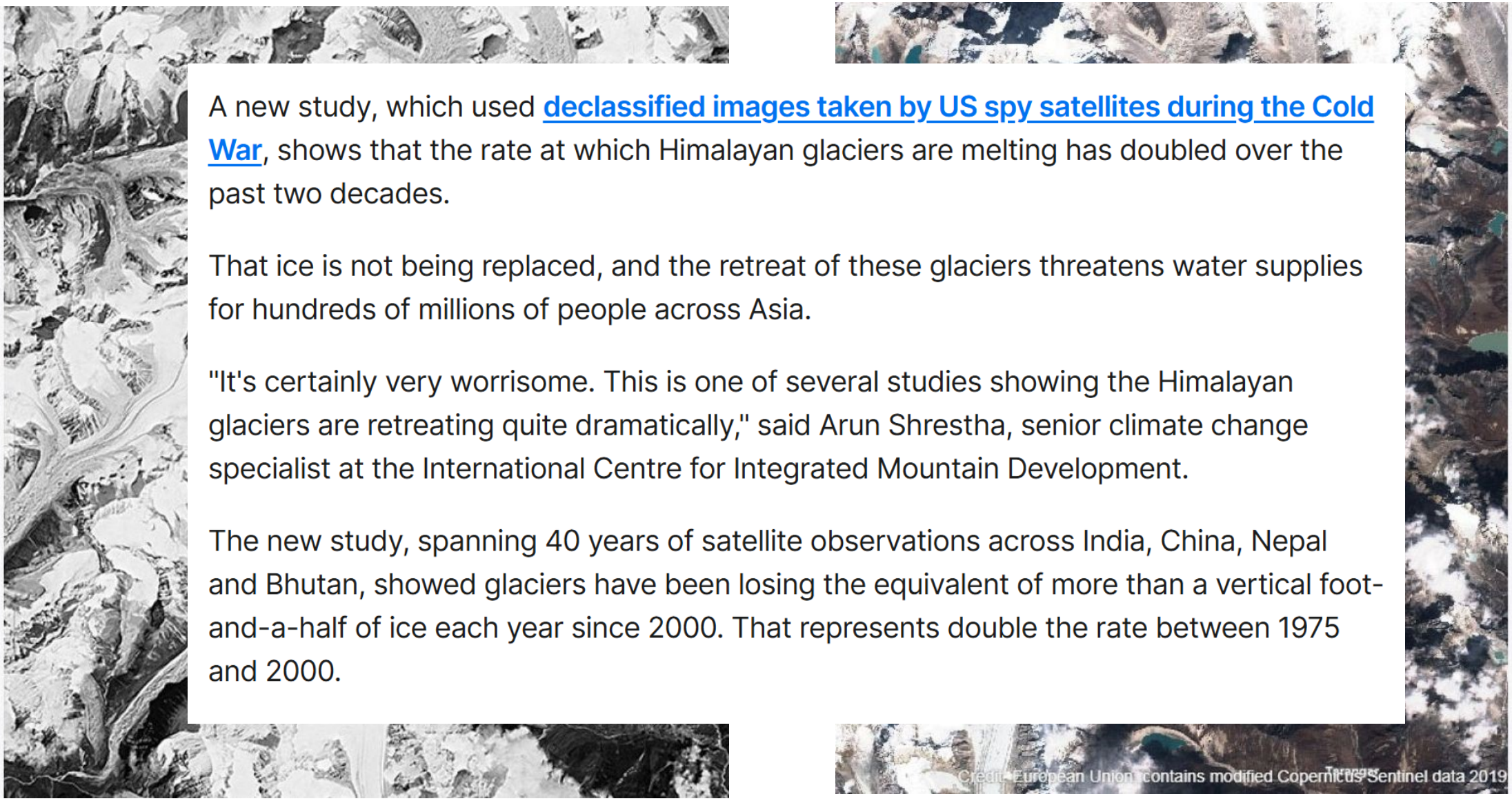
Low-res is free

Daily to monthly updates

Optical, infrared, radar, ...

Archival data ~20 years

Melting Glaciers



A new study, which used declassified images taken by US spy satellites during the Cold War, shows that the rate at which Himalayan glaciers are melting has doubled over the past two decades.

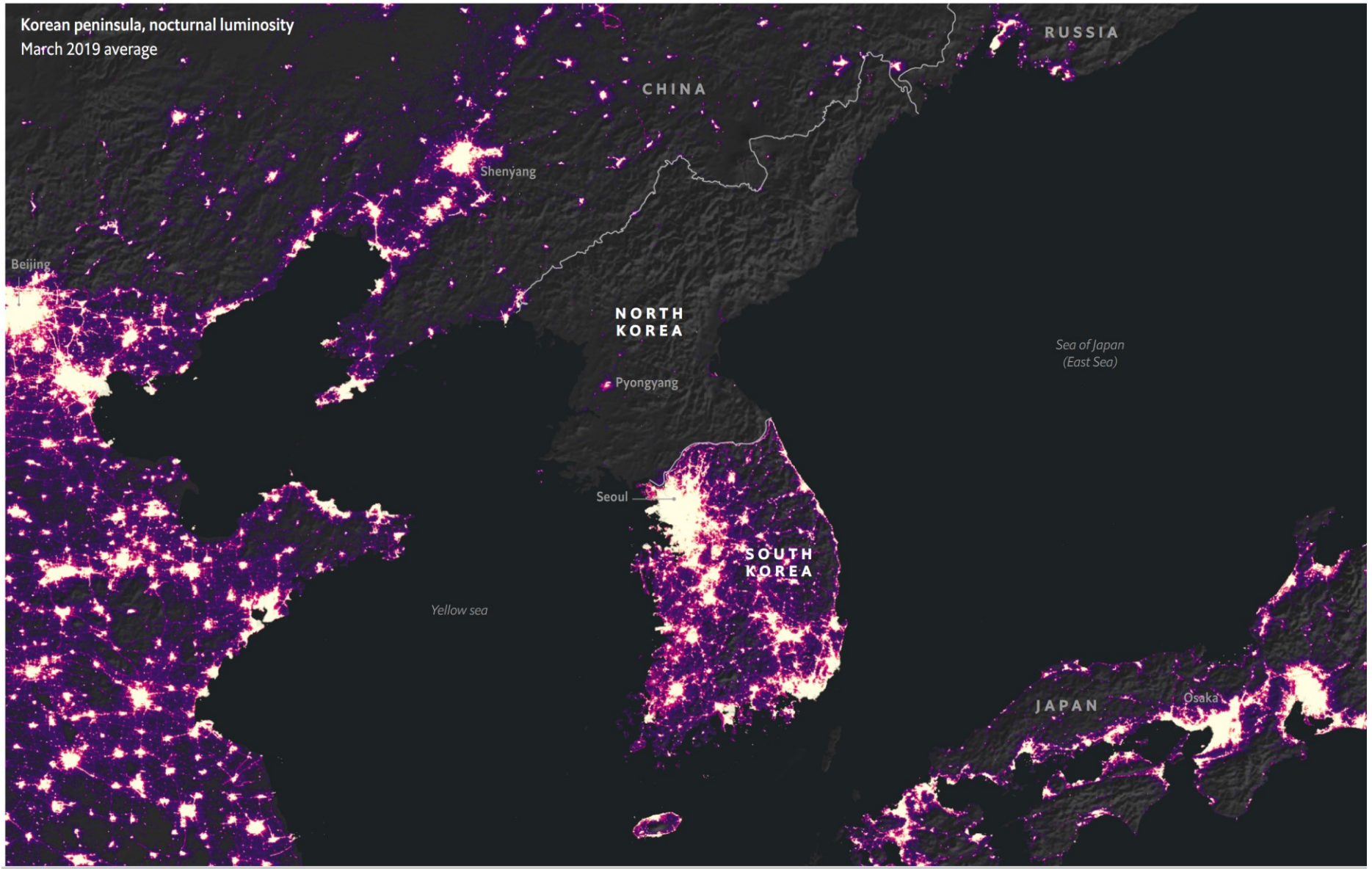
That ice is not being replaced, and the retreat of these glaciers threatens water supplies for hundreds of millions of people across Asia.

"It's certainly very worrisome. This is one of several studies showing the Himalayan glaciers are retreating quite dramatically," said Arun Shrestha, senior climate change specialist at the International Centre for Integrated Mountain Development.

The new study, spanning 40 years of satellite observations across India, China, Nepal and Bhutan, showed glaciers have been losing the equivalent of more than a vertical foot-and-a-half of ice each year since 2000. That represents double the rate between 1975 and 2000.

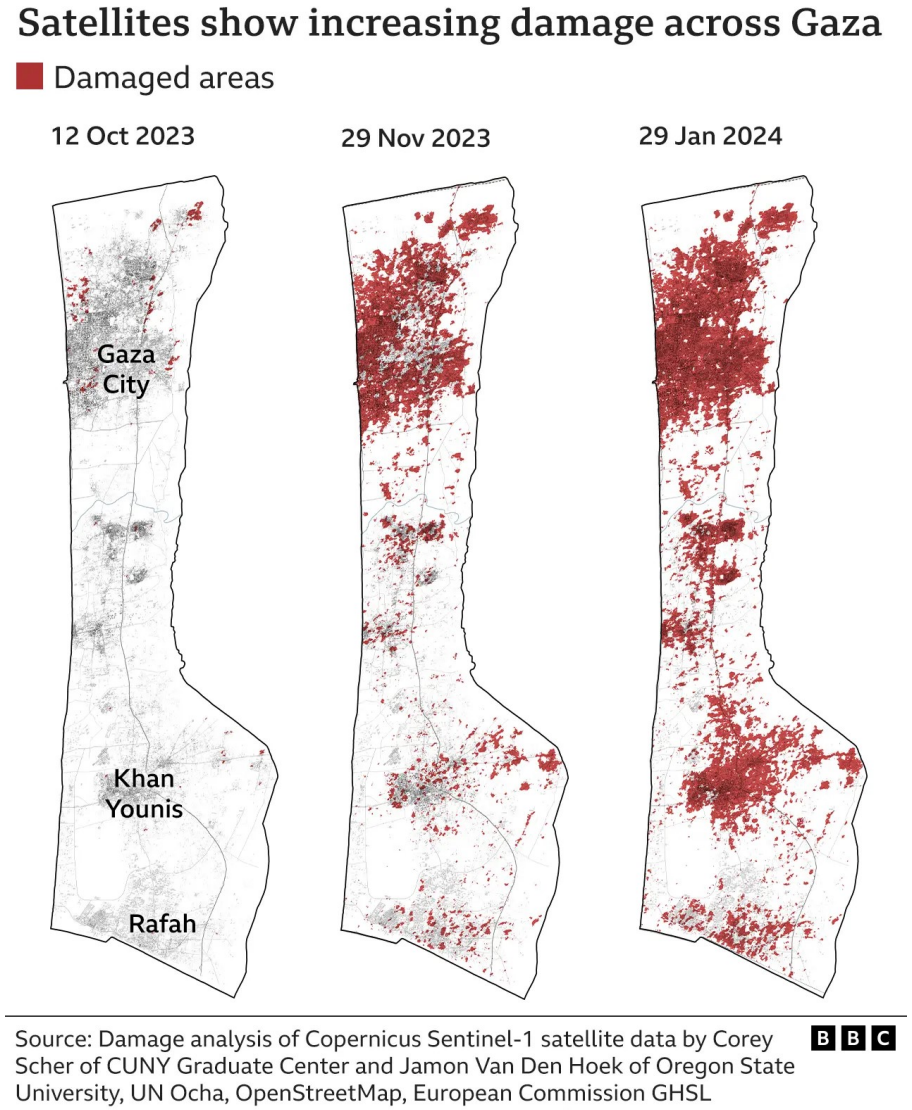
https://www.euronews.com/2019/06/20/himalayan-glaciers-melting-at-alarming-rate-satellite-data-shows

Night Light in North Korea



https://www.economist.com/graphic-detail/2019/05/04/satellite-data-shed-new-light-on-north-koreas-opaque-economy

Destruction in the Gaza War



https://www.bbc.com/news/world-middle-east-68006607

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

SIC Saarland Informatics Campus

Mapping Wealth and Poverty

Detecting Tax Evasion from Space

SPIEGEL International

Finding Swimming Pools with Google Earth

Greek Government Hauls in Billions in Back Taxes

In a bid to increase revenues, the Greek authorities are employing all kinds of clever tricks to crack down on tax cheats, including using Google Earth to find undeclared swimming pools. But efforts by the government to liberalize markets could unleash a wave of civil unrest.

Von **Daniel Steinvorth**
02.08.2010, 14.48 Uhr

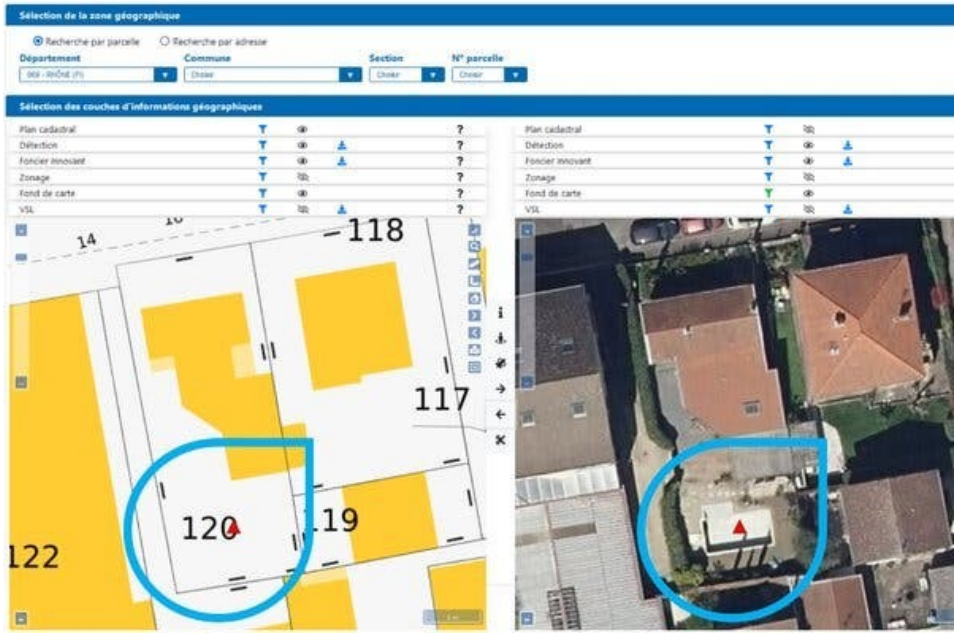
"They use satellite pictures by Google Earth to locate country villas, swimming pools and properties. And these tactics have revealed that the suburbs didn't have 324 swimming pools, as was reported, but rather 16,974."

https://www.spiegel.de/international/europe/finding-swimming-pools-with-google-earth-greek-government-hauls-in-billions-in-back-taxes-a-709703.html

The New York Times

French Tax Collectors Use A.I. to Spot Thousands of Undeclared Pools

Algorithms combing through satellite photos found over 20,000 unreported swimming pools in a few regions, yielding an expected $10 million in taxes, and the system will soon go nationwide.



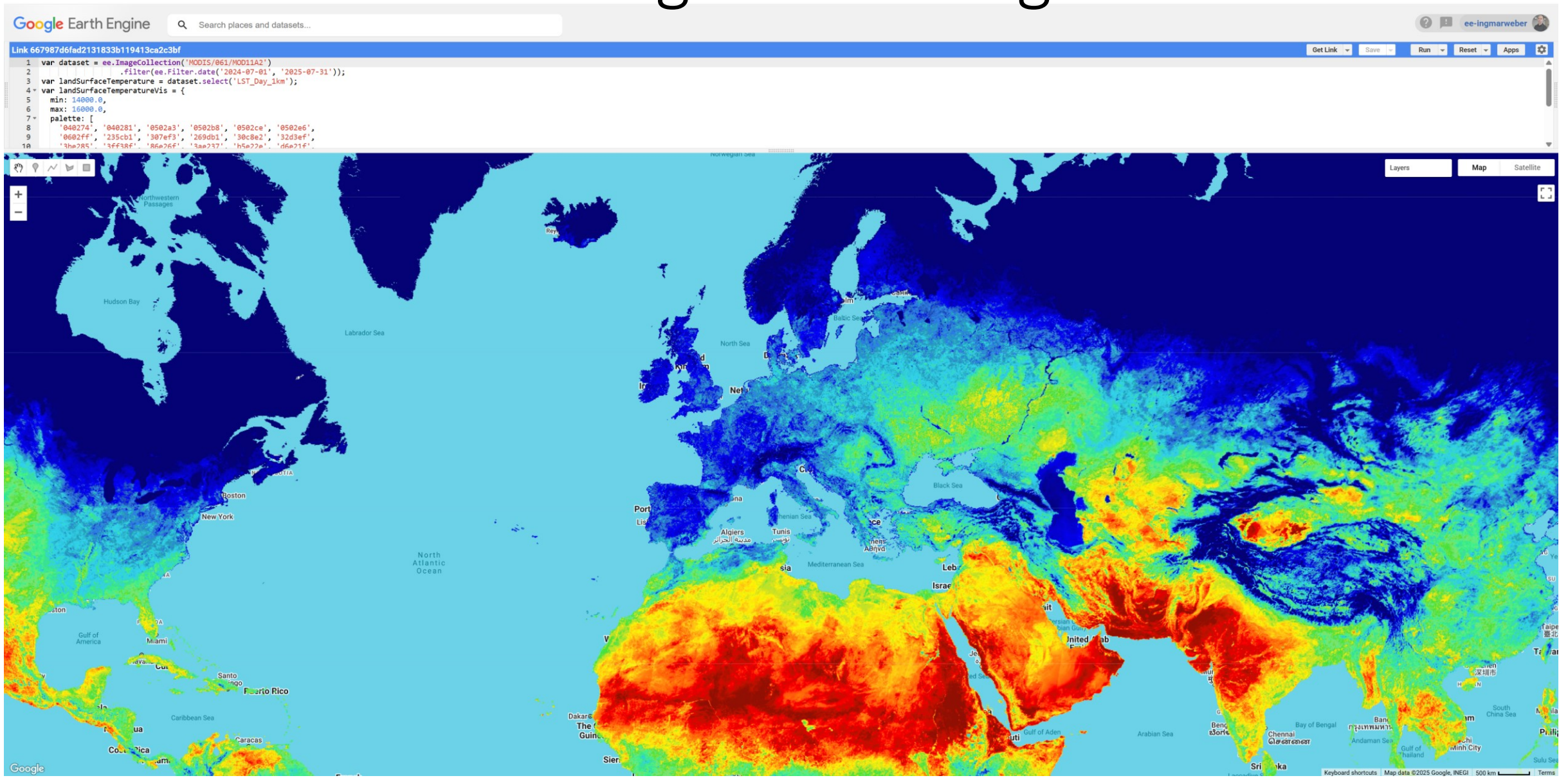
https://www.nytimes.com/2022/08/30/world/europe/france-taxes-pools-artificial-intelligence.html

Racism and Climate Impact

In cities like Baltimore, Dallas, Denver, Miami, Portland and New York, neighborhoods that are poorer and have more residents of color can be 5°-20° Fahrenheit (3°-5°C) hotter in summer than wealthier, whiter parts of the same city.

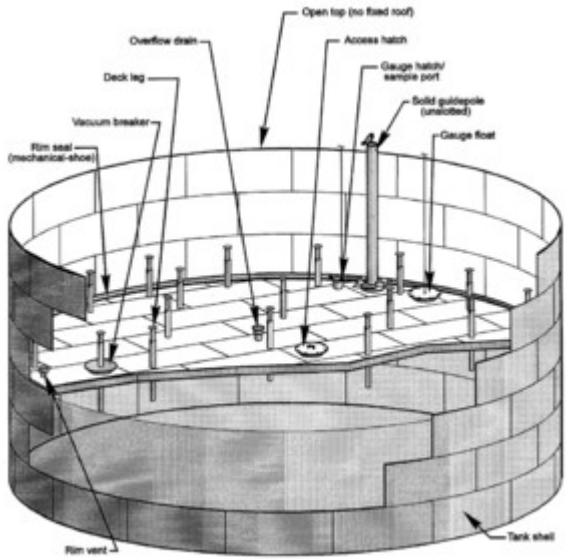
https://www.nytimes.com/interactive/2020/08/24/climate/racism-redlining-cities-global-warming.html

Data via Google Earth Engine



https://code.earthengine.google.com/667987d6fad2131833b119413ca2c3bf

Floating Roof Oil Tanks



https://en.wikipedia.org/wiki/External_floating_roof_tank

https://www.youtube.com/watch?v=aKIpZU95jco

https://medium.com/from-the-macroscope/as-world-panics-over-oil-storage-orbital-data-shows-2b-barrels-of-space-left-d6745216e946

Not Just Daytime …

Can You See Displacement from Space?

https://www.nytimes.com/interactive/2020/02/14/world/middleeast/syria-idlib-displaced.html

https://www.aljazeera.com/gallery/2020/2/12/in-pictures-syrians-flee-for-safety-as-battle-in-northwest-rages/

Two Approaches to Monitoring Displacement

Vehicle Flows
Look for cars going from A to B
Captures movement in the moment

Vehicle Stocks
Two snapshots: difference in vehicle counts
Captures before-after differences

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

Cars Leaving Parts of Ukraine

SIC Saarland Informatics Campus

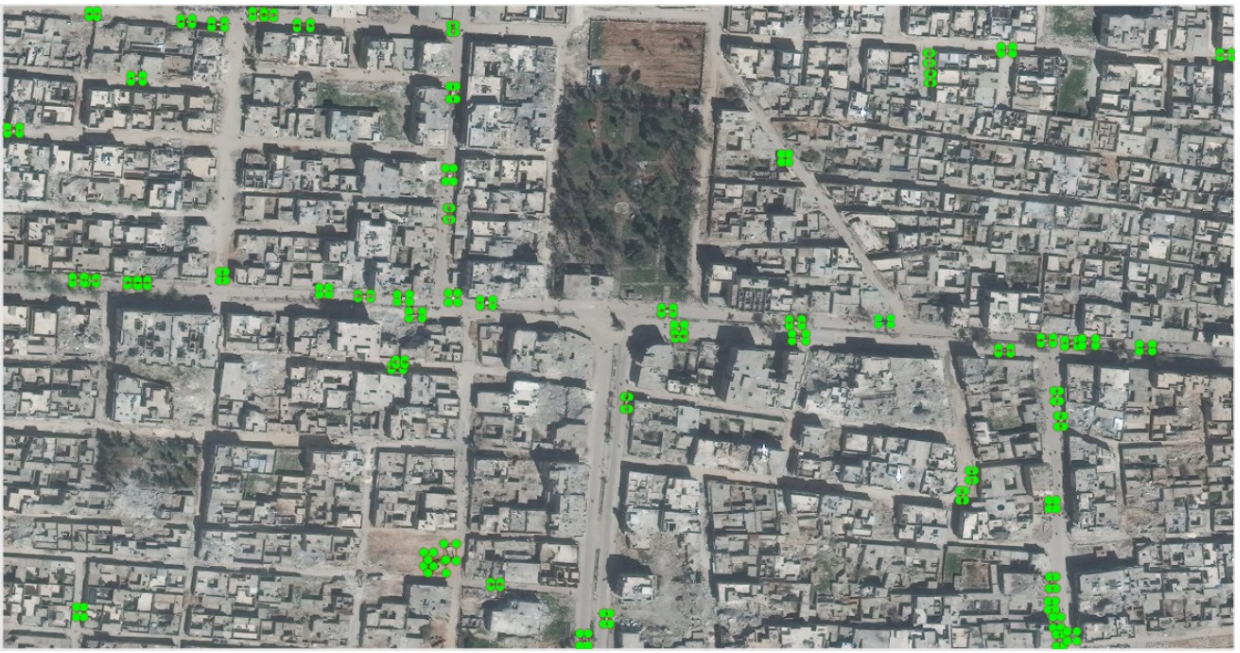
Kharkiv

Before invasion

After invasion

Different lighting, different nadir angle, different day of week, …

Works Well for High-Res Images



Manually annotated vehicles

Automatically identified vehicles

Manually annotated vehicles: **56**

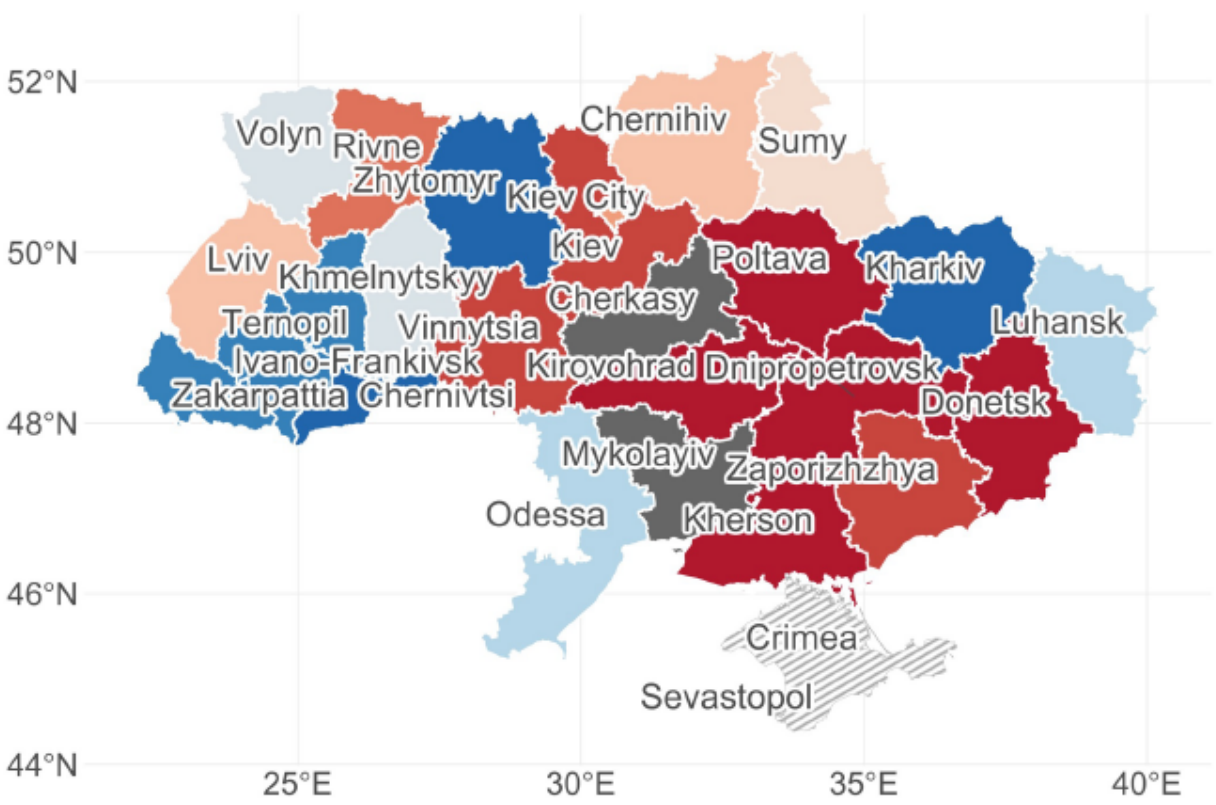
Automatically identified vehicles: **46 / 56** plus **2** false positives

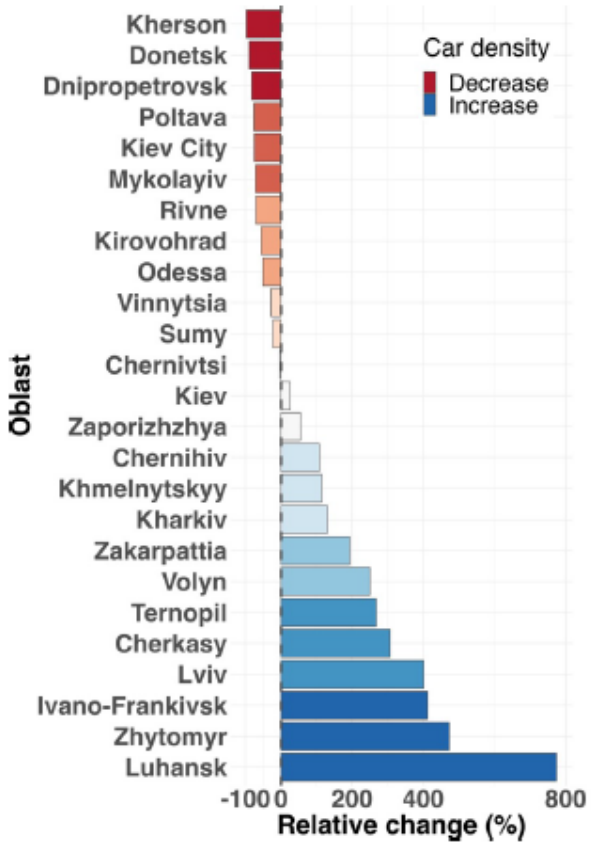
Precision: 46 / 48 = .96

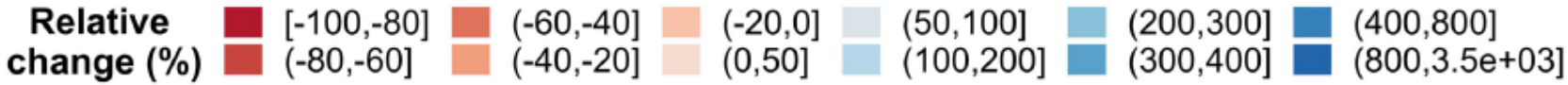
Recall: 46 / 56 = .82

… using off-the-shelf satellite object detection libraries

Shifts in car density 2019 -> Post-Invasion 2022



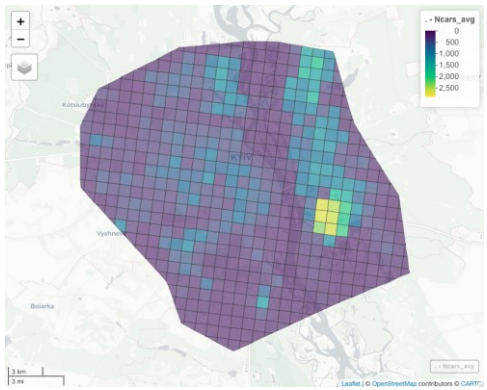




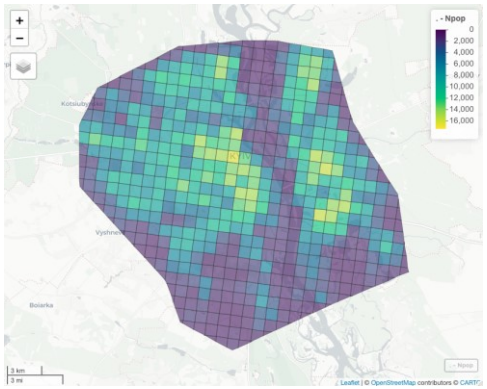
From Cars to Population Estimates

Historic data for calibration

cars vs. gridded population



Pre-War # Cars



Pre-War Population

Global

Learn a *city-wide* correspondence:
Any grid cell in Kiev with x cars corresponds to y people
+ Can learn that x=0, but y!=0
+ Can fit any functional shape
- Fooled by local anomalies, such as industrial areas

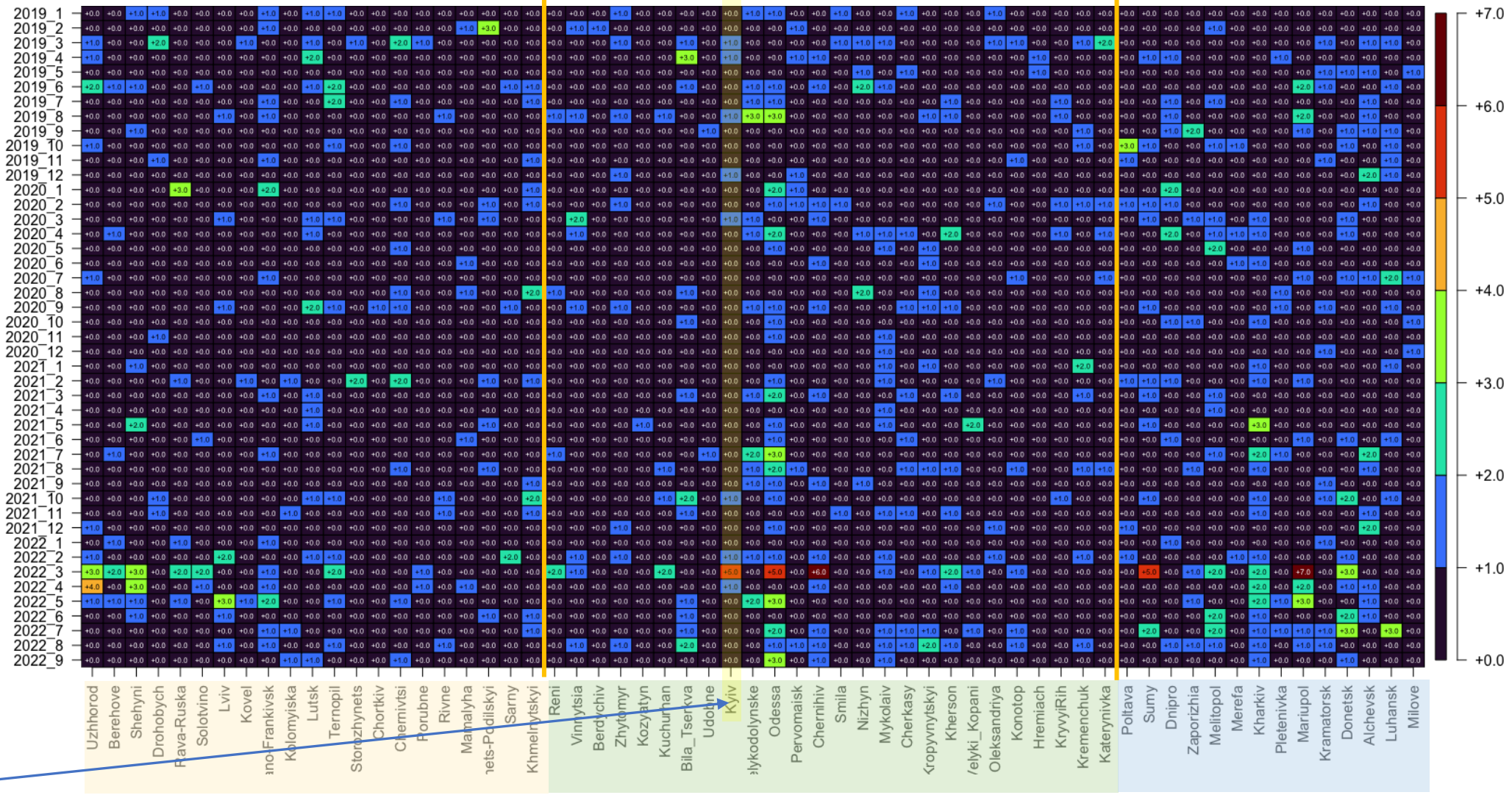
Local

Learn a *per-grid-cell* correspondence
For grid cell i, x_i cars corresponds to y_i people
+ Naturally adapts to local anomalies
- Assumes a x_i/2 => y_i/2

Very simple yet effective. Could incorporate many more covariates:
- Public transportation, roads, population density, (street) parking availability, …

Sort of Works, But Data Availability is a Problem…

MAXAR



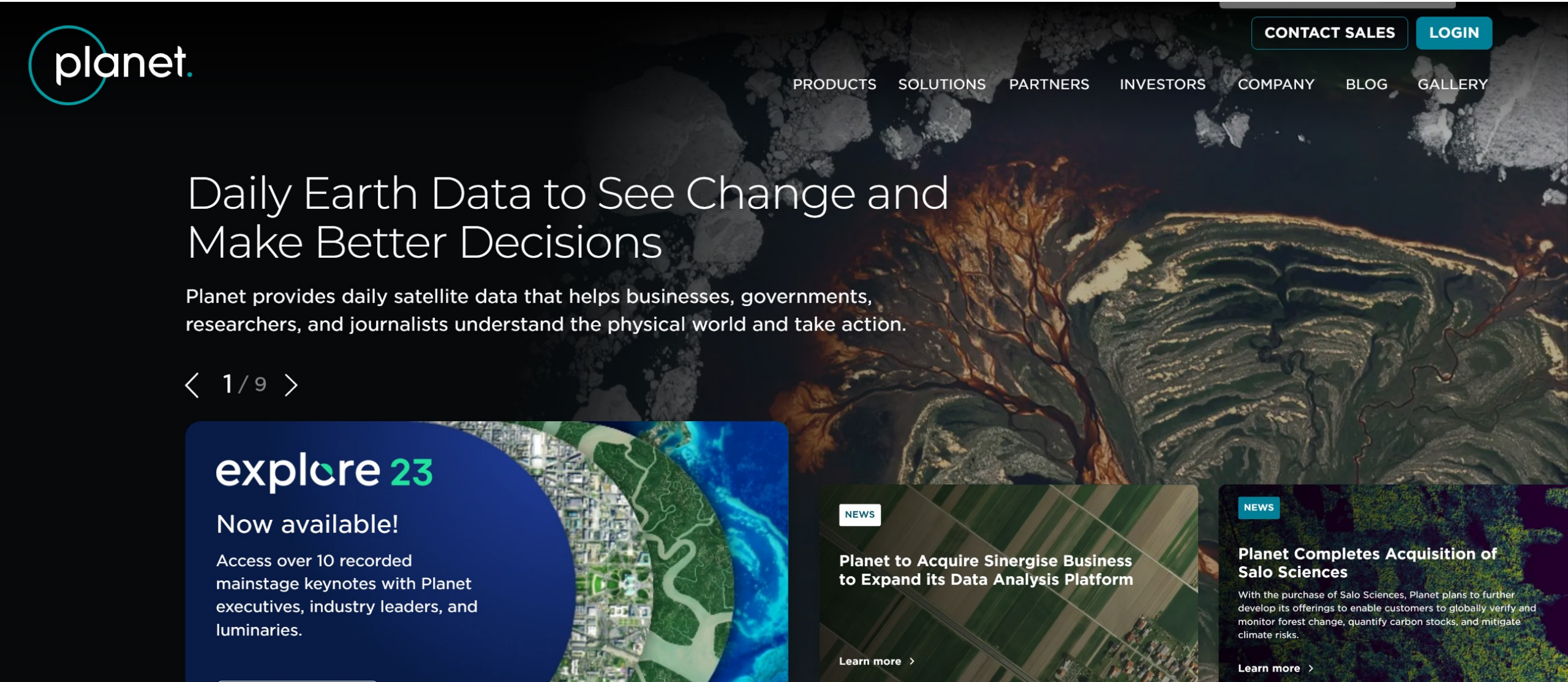
Most months without *any* data for Kyiv

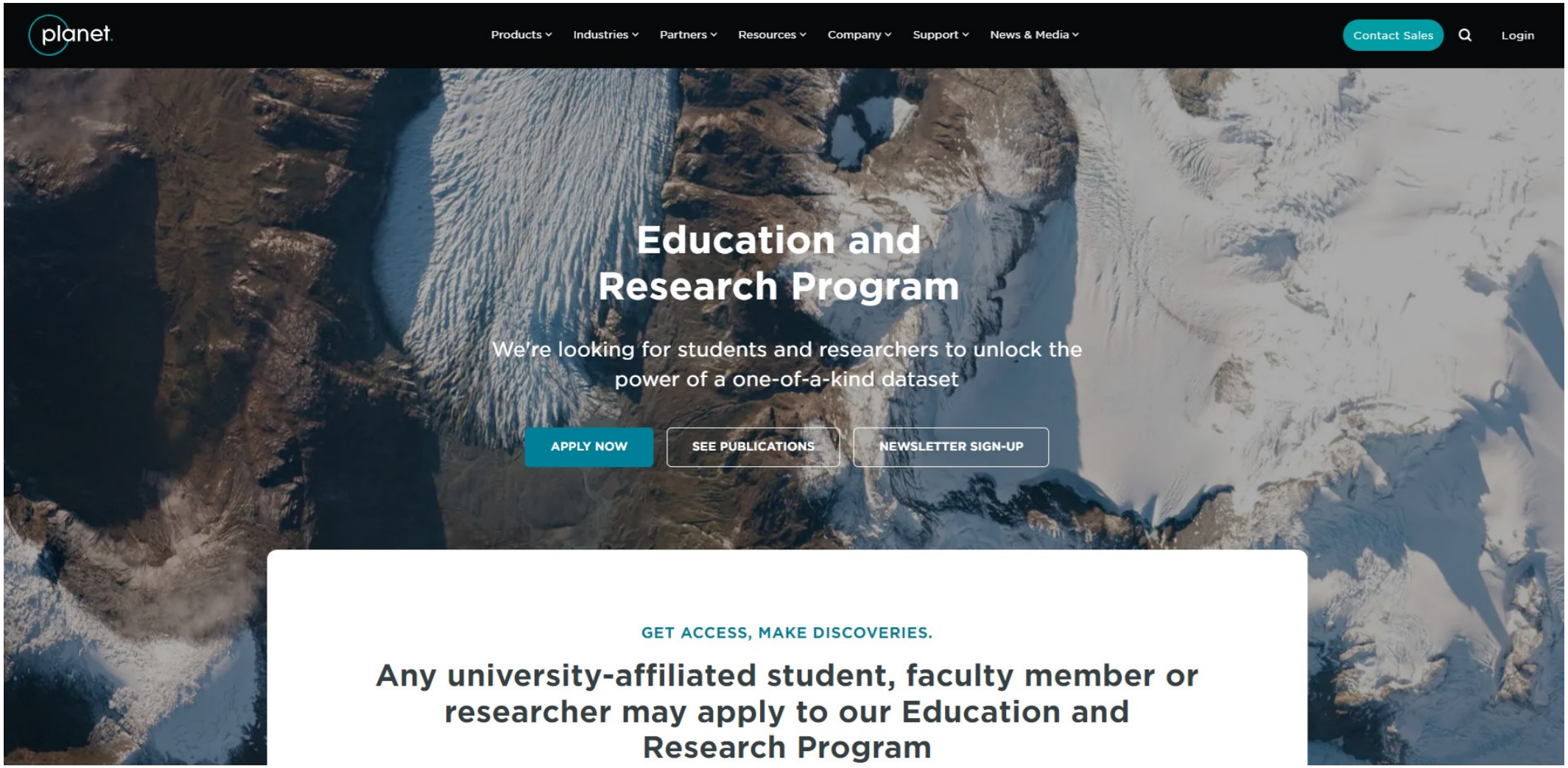
West

Central

East

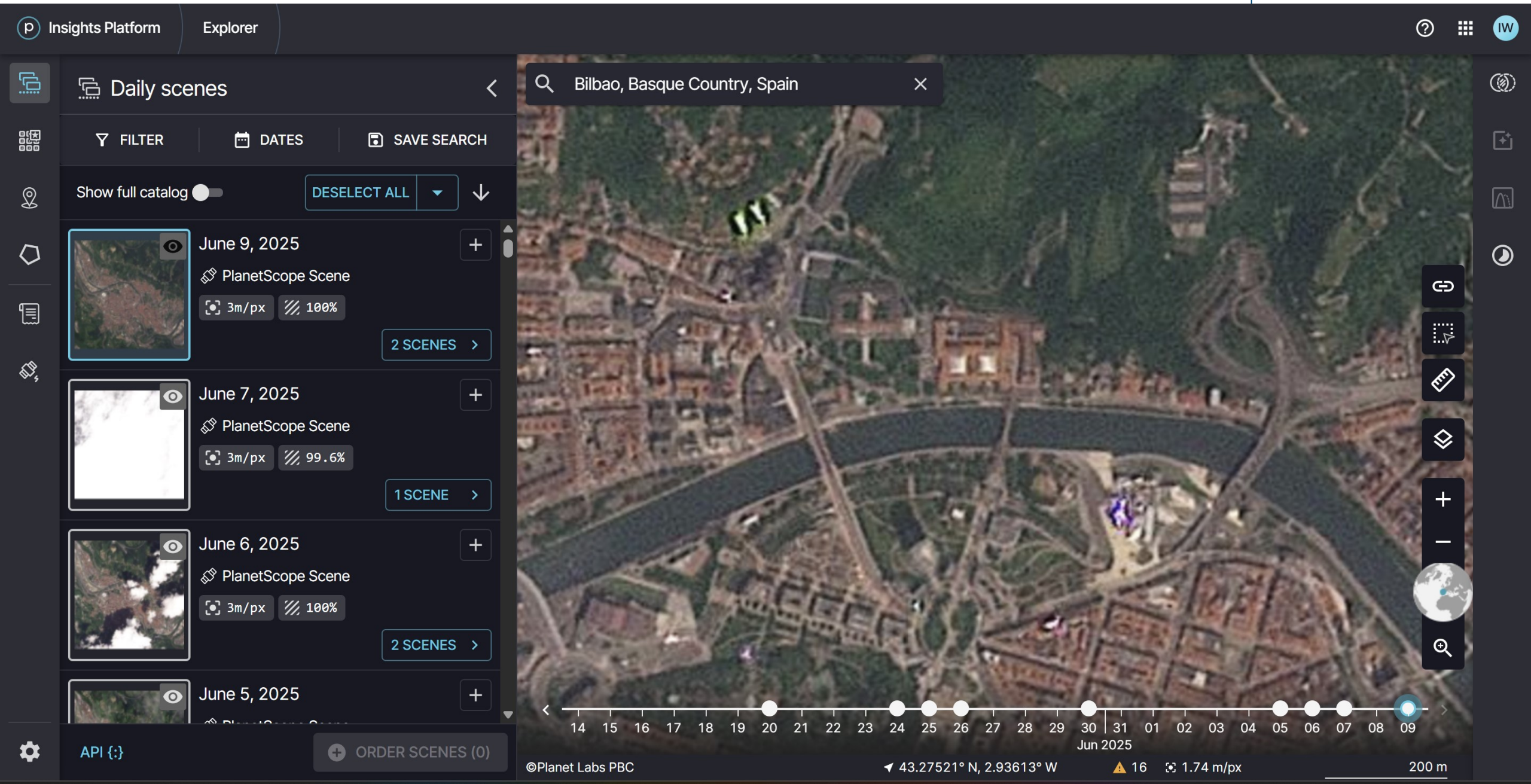
Can We Use *daily* planet.com Imagery Instead?





https://www.planet.com/markets/education-and-research/

Pixels are Bigger than Individual Cars



How can you count cars, if you can’t even see them?

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

Day of the Week as Weak Annotation

SIC Saarland Informatics Campus

The model learns to estimate relative car park occupancy.

Given a pair of images, which one is fuller?

Image 1

Same parking lot

Image 2

Parking lot size	AUC
large parking lots	**0.92**
medium parking lots	0.91
small parking lots	0.65

Performance on 20% of unseen parking lots

Why do we care about parking lots?

- Monitor (internal) displacement

Are cars leaving the cities? Are border cities filling up?

- Monitor for new epidemics?

Are parkings lots of hospitals and clinics fuller than usual?

- Monitor economic activity?

Are companies reducing their staff? Are people (not) going shopping?

… in regions where it’s not always cloudy and where people move around by car.

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

SIC Saarland Informatics Campus

Computational Social Science from Space

Topic: The IC2S2 community studies how online social media and web data can be used to understand different aspects of society and human behavior. While these online data sources are valuable for sensing and quantifying the 'social fabric', they are not made for sensing the physical world. This tutorial will provide attendees with another data source with which to complement their analysis: satellite imagery. As the Earth Observation (EO) community is turning to use social media data, we aim to encourage a reciprocal interdisciplinary approach, too. At the same time, satellite-based remote sensing creates new needs for setting community standards around the responsible use of such data, something the IC2S2 community is well-positioned to contribute to. In this tutorial, we will give a short overview of existing work in the social sciences that uses EO to study social phenomena, encouraging discussion from the audience. Additionally, we will offer a hands-on introduction to using Google Earth Engine to demonstrate the effects of redlining, the practice of racially segregated housing policies, visible from space and still impacting communities today.

https://sites.google.com/view/cssfromspace

Advertising Data

Social Media Data in Post-API Age

THE CONVERSATION

Rigor académico, oficio periodístico

Buscar análisis, investigaciones...

Ciencia + Tecnología | Cultura | Economía | Educación | Salud | Medioambiente | Política + Sociedad | Júnior | Bienestar digital y menores

Facebook's data lockdown is a disaster for academic researchers

Publicado: 11 abril 2018 12:10 CEST

Shutterstock

Copy link
Correo
X (Twitter)
Bluesky

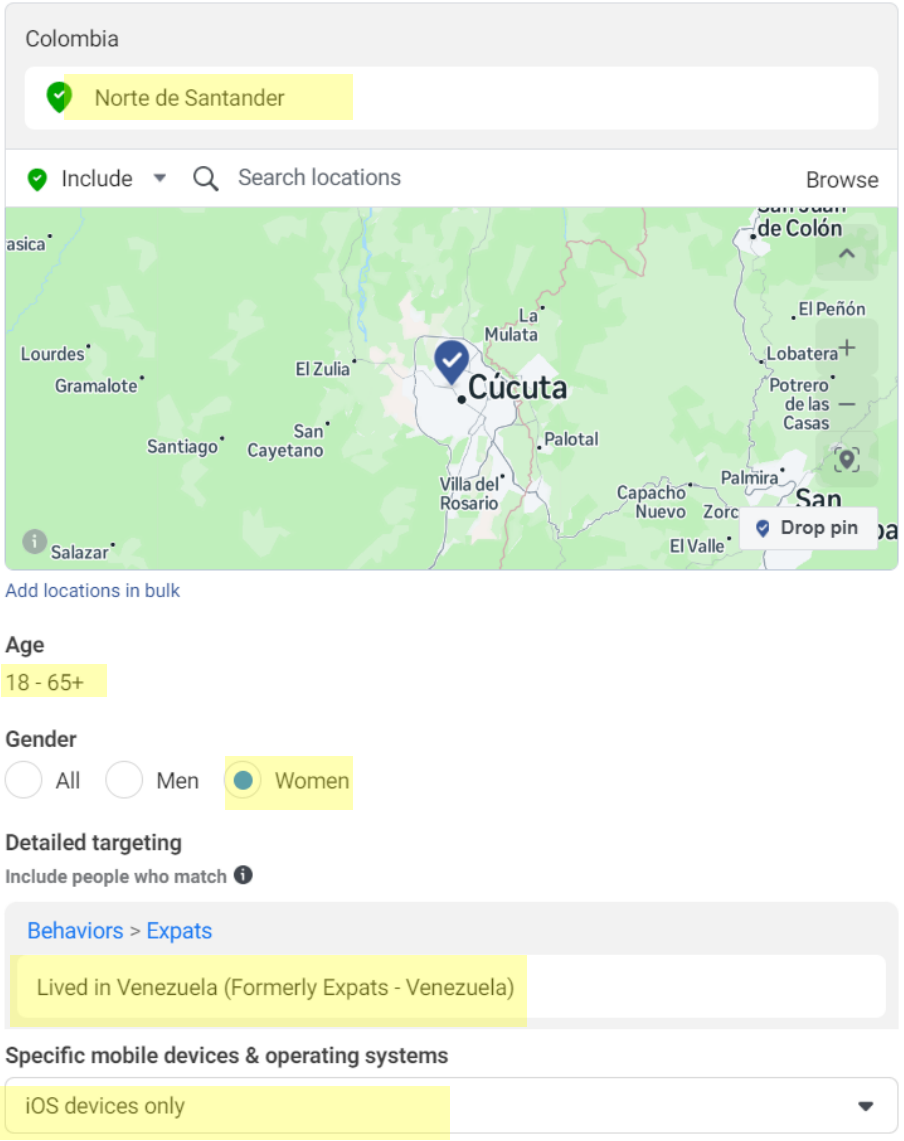
Facebook recently announced dramatic data access restrictions on its app and website. The company framed the lockdown as an attempt to protect user information, in response to the public outcry following the Cambridge Analytica scandal.

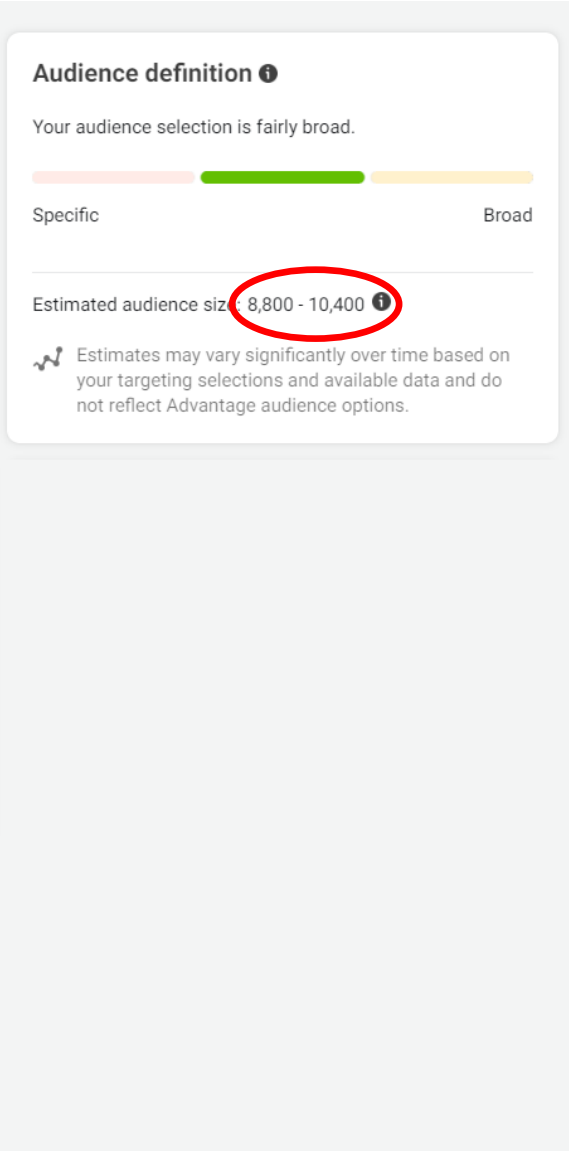
Autoría

Marco Bastos
Senior Lecturer in Media and Communication, City St George's, University of London

https://theconversation.com/facebooks-data-lockdown-is-a-disaster-for-academic-researchers-94533

Tapping into Surveillance Capitalism





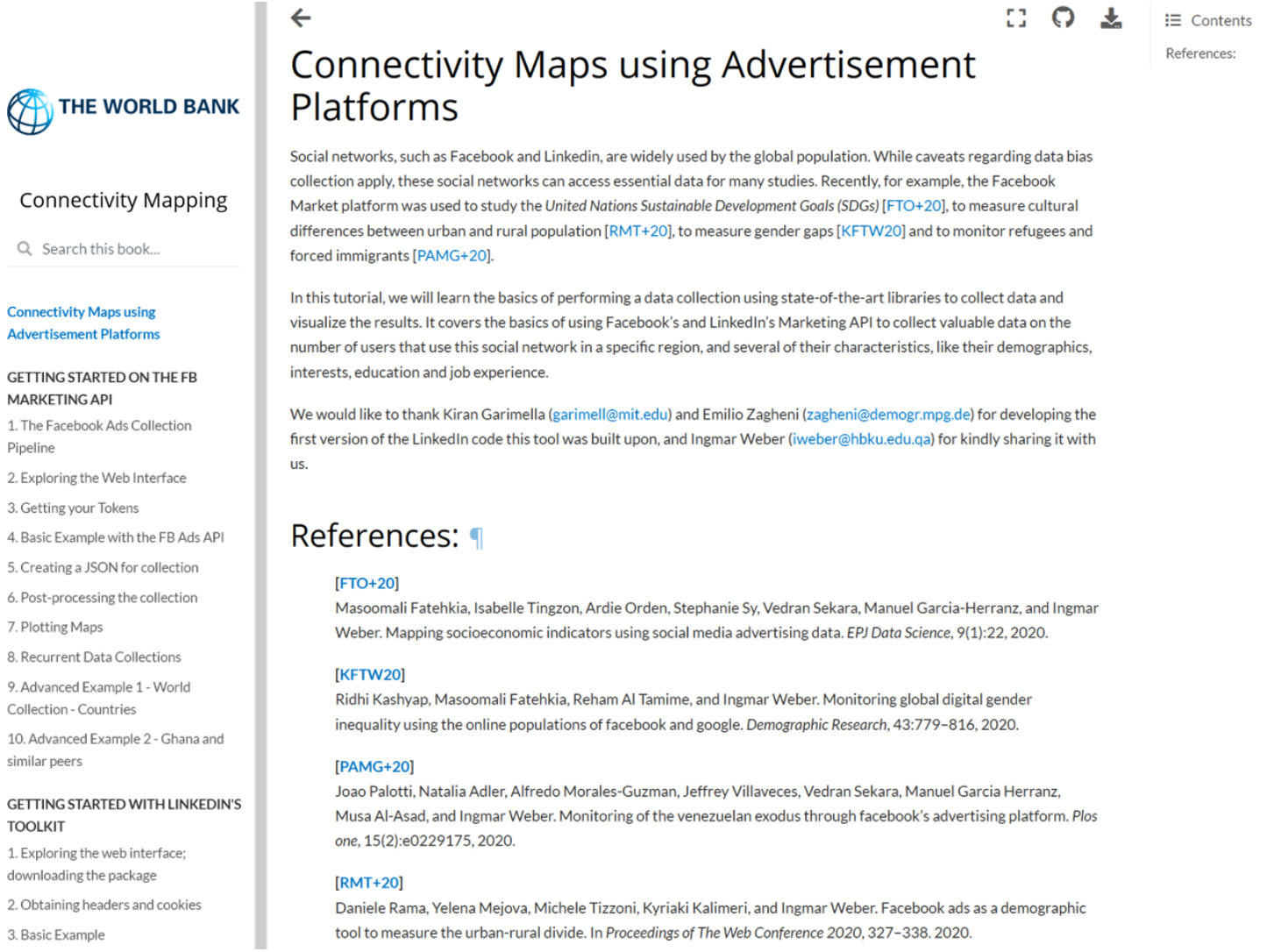
Real-time "census" across 3 billion Facebook (+ Instagram + Messenger) users

Provides: How many users match criteria X?

No individual data

Available for free via API

Well-Documented API and Wrapper



THE WORLD BANK

Connectivity Mapping

Search this book...

Connectivity Maps using Advertisement Platforms

GETTING STARTED ON THE FB MARKETING API

1. The Facebook Ads Collection Pipeline
2. Exploring the Web Interface
3. Getting your Tokens
4. Basic Example with the FB Ads API
5. Creating a JSON for collection
6. Post-processing the collection
7. Plotting Maps
8. Recurrent Data Collections
9. Advanced Example 1 - World Collection - Countries
10. Advanced Example 2 - Ghana and similar peers

GETTING STARTED WITH LINKEDIN'S TOOLKIT

1. Exploring the web interface; downloading the package
2. Obtaining headers and cookies
3. Basic Example

Contents

References:

Connectivity Maps using Advertisement Platforms

Social networks, such as Facebook and Linkedin, are widely used by the global population. While caveats regarding data bias collection apply, these social networks can access essential data for many studies. Recently, for example, the Facebook Market platform was used to study the *United Nations Sustainable Development Goals (SDGs)* [FTO+20], to measure cultural differences between urban and rural population [RMT+20], to measure gender gaps [KFTW20] and to monitor refugees and forced immigrants [PAMG+20].

In this tutorial, we will learn the basics of performing a data collection using state-of-the-art libraries to collect data and visualize the results. It covers the basics of using Facebook's and LinkedIn's Marketing API to collect valuable data on the number of users that use this social network in a specific region, and several of their characteristics, like their demographics, interests, education and job experience.

We would like to thank Kiran Garimella (garimell@mit.edu) and Emilio Zagheni (zagheni@demogr.mpg.de) for developing the first version of the LinkedIn code this tool was built upon, and Ingmar Weber (iweber@hbku.edu.qa) for kindly sharing it with us.

References: ¶

[FTO+20]
Masoomali Fatehkia, Isabelle Tingzon, Ardie Orden, Stephanie Sy, Vedran Sekara, Manuel Garcia-Herranz, and Ingmar Weber. Mapping socioeconomic indicators using social media advertising data. *EPJ Data Science*, 9(1):22, 2020.

[KFTW20]
Ridhi Kashyap, Masoomali Fatehkia, Reham Al Tamime, and Ingmar Weber. Monitoring global digital gender inequality using the online populations of facebook and google. *Demographic Research*, 43:779–816, 2020.

[PAMG+20]
Joao Palotti, Natalia Adler, Alfredo Morales-Guzman, Jeffrey Villaveces, Vedran Sekara, Manuel Garcia Herranz, Musa Al-Asad, and Ingmar Weber. Monitoring of the venezuelan exodus through facebook's advertising platform. *Plos one*, 15(2):e0229175, 2020.

[RMT+20]
Daniele Rama, Yelena Mejova, Michele Tizzoni, Kyriaki Kalimeri, and Ingmar Weber. Facebook ads as a demographic tool to measure the urban-rural divide. In *Proceedings of The Web Conference 2020*, 327–338. 2020.

https://worldbank.github.io/connectivity_mapping/intro.html

Venezuelan Crisis and the “Maduro Diet”

Inflation >130,000% (2018, Central Bank)

Unemployment > 35% (2018, IMF)

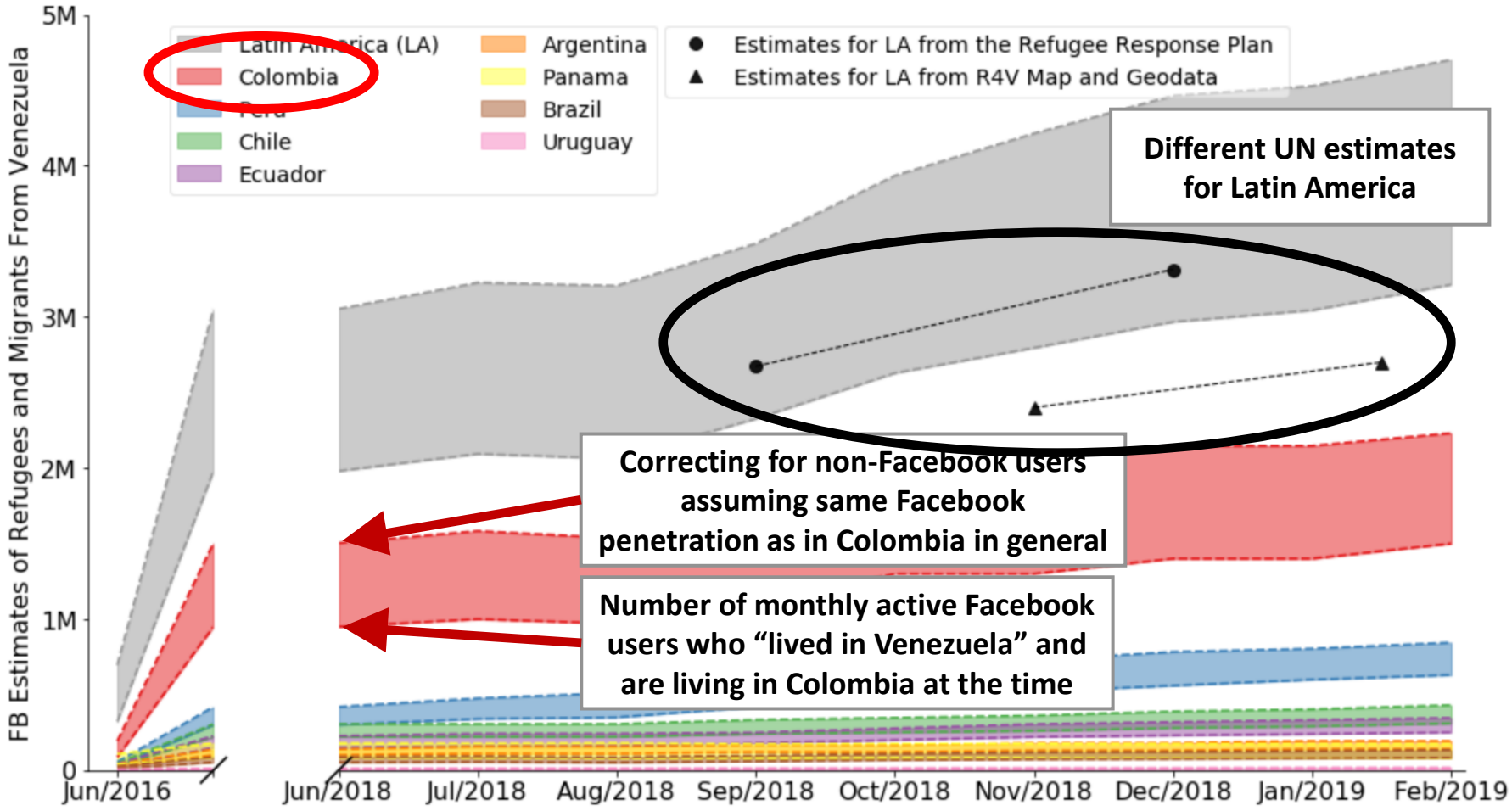
Minimum wage < 1000kcal/day (2018, NPR)

Foto AP, 2016
https://www.americasquarterly.org/article/the-maduro-diet-a-photo-essay-from-venezuela/

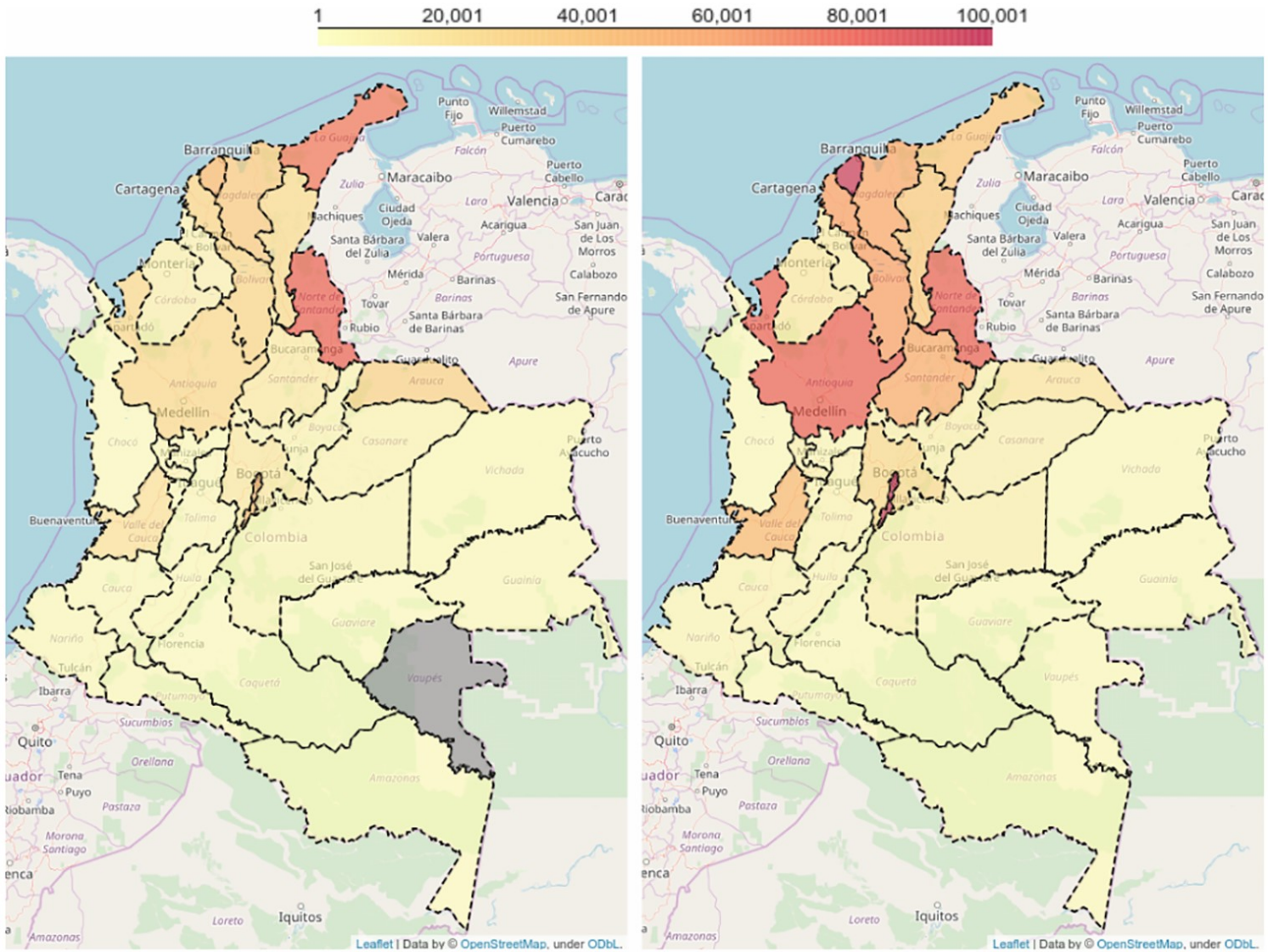
Migration and Refugee Crisis

Simon Bolivar International Bridge in February 2018 (EPA/Rex/Shutterstock)
https://www.bbc.co.uk/news/resources/idt-sh/Venezuela_bridge

Monitoring Trends in Real-Time



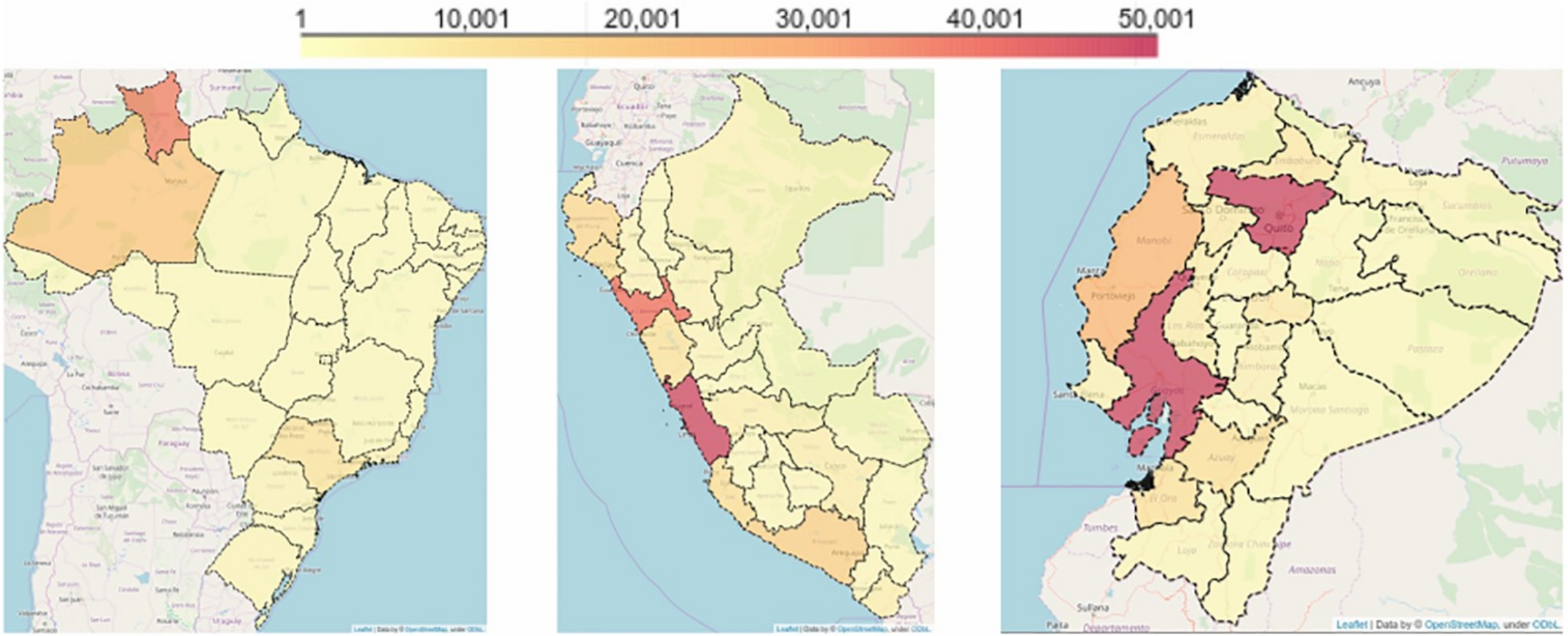
Validation w/ (Few) Available Data



(a) RAMV - April 6 to June 8, 2018

(b) Facebook - June 8, 2018

Previously Unavailable Estimates

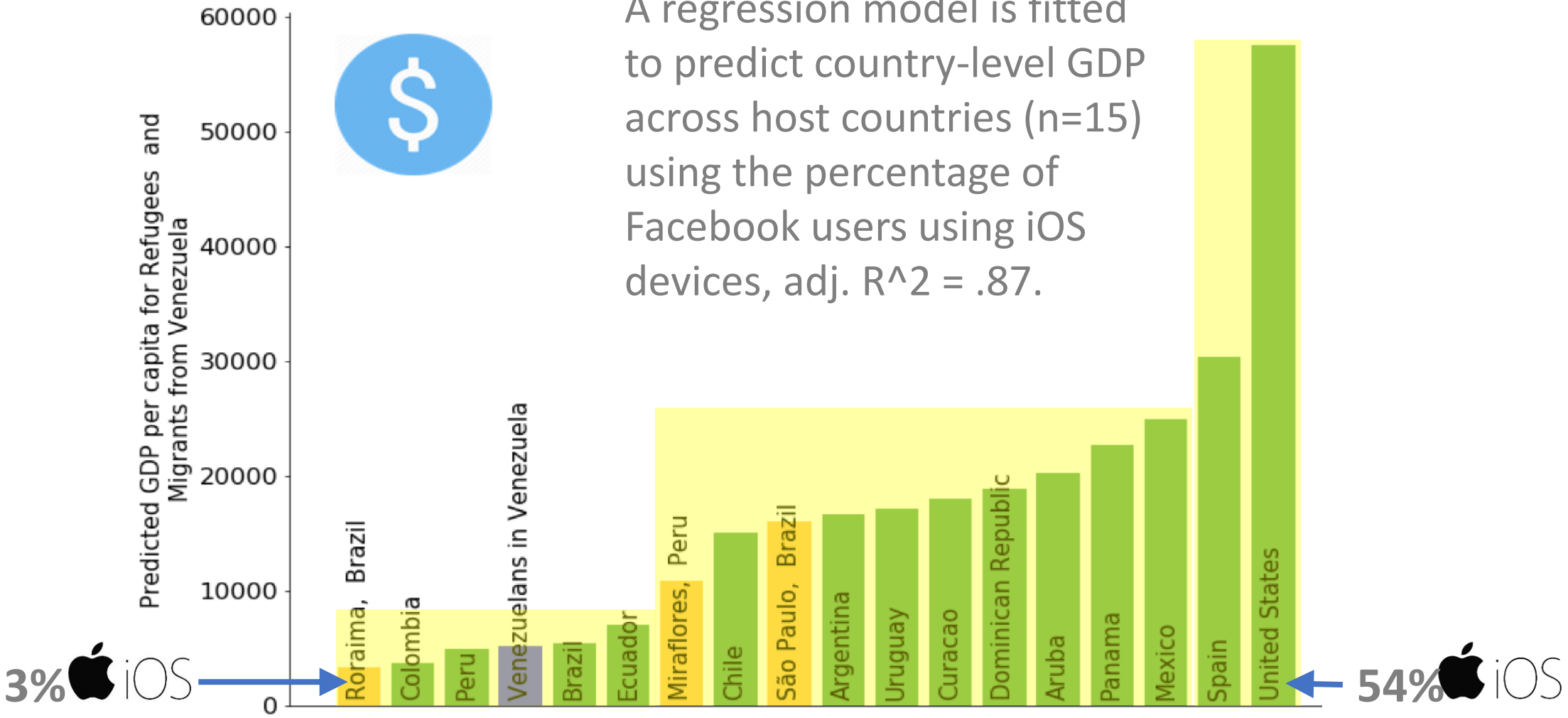


(a) Brazil (b) Peru (c) Ecuador

Facebook. Feb 2019

Inferred Vulnerability Based on OS

A regression model is fitted to predict country-level GDP across host countries (n=15) using the percentage of Facebook users using iOS devices, adj. R^2 = .87.



Operational Impact

"We relied on their [QCRI's] research analysis for the Venezuela situation. Understanding population movement and density allows humanitarian agencies, like ours, to provide humanitarian assistance and protection in specific areas where population is more concentrated."

Rebeca Moreno Jimenez
UNHCR Lead Data Scientist and Innovation Officer

"Knowing this distribution helped redefine and amplify the geographical scope […]. The recognition of the national scale of the crisis was particularly relevant for UNICEF to develop anti-xenophobia campaigns using Facebook's chatbots."

Natalia Adler
Former UNICEF Data, Research, Policy Manager

Continued Operational Use

iMMAP | Inicio | Contexto | Sectores Priorizados | Proyectos | Productos de Información | Publicaciones | Contáctenos

LOS RESULTADOS

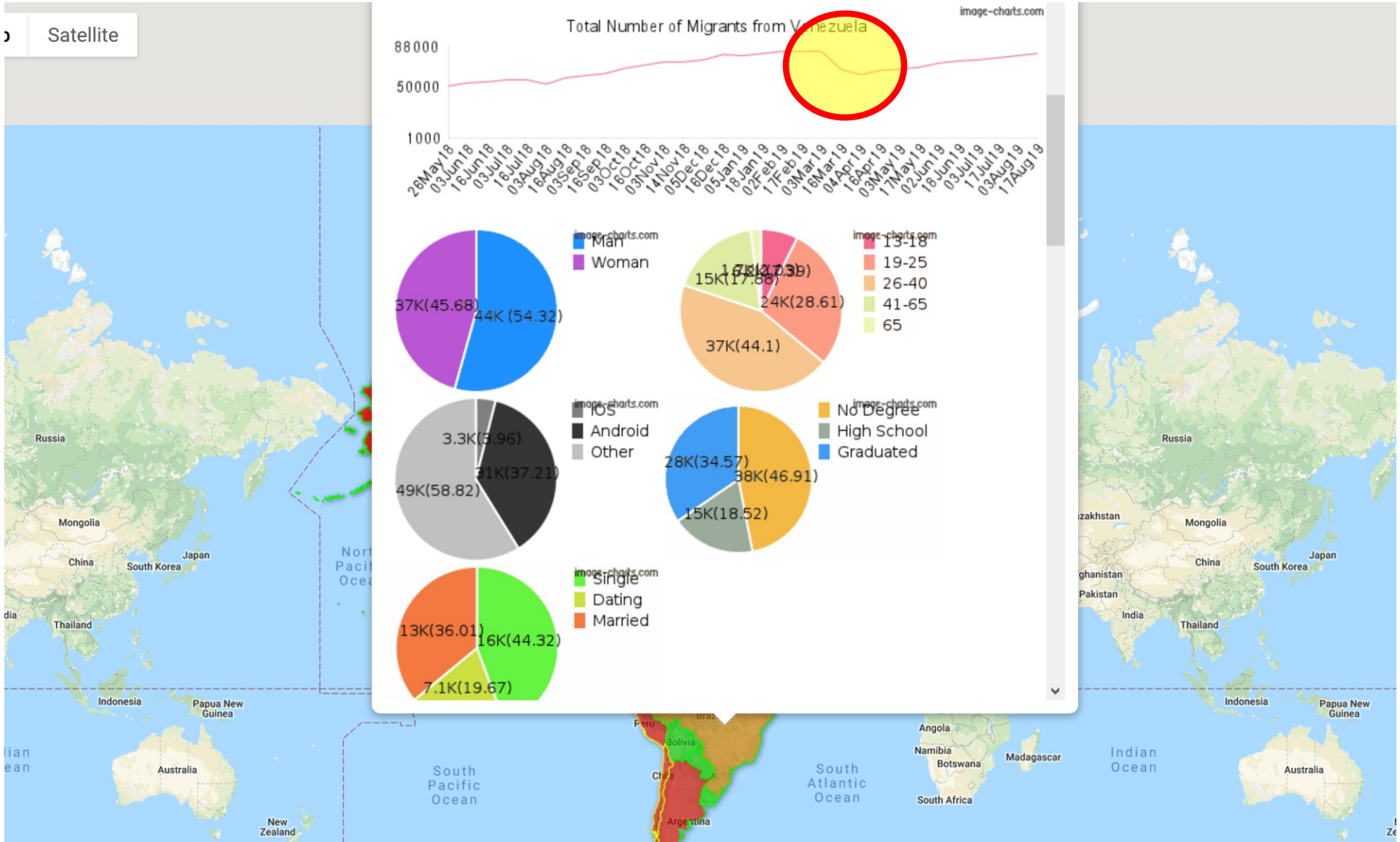
USAID | iMMAP | Detección de Usuarios venezolanos conectados en Facebook | Municipal | Departamental | Regional | Perú | Ecuador | Tendencia

IMMAP localiza a los migrantes venezolanos en América Latina a través del uso de el api de Facebook advertising data mostrando las conexiones de usuarios que antes vivian en Venezuela y ahora viven en el extranjero.

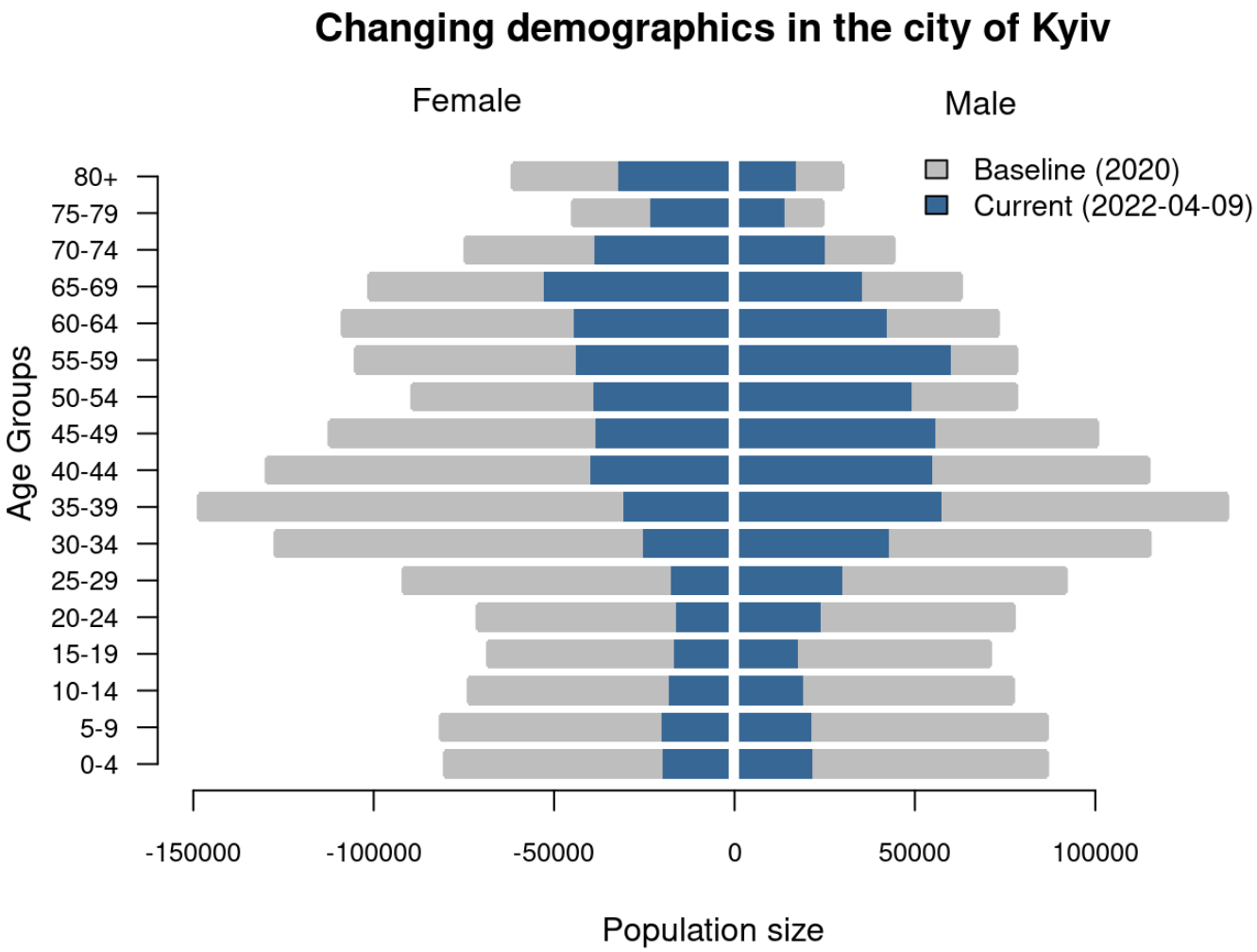
Total usuarios conectados	Dato Oficial Migración Colombia
1,845,200 Usuarios conectados Febrero 15	**1,729,537** Venezolanos en Colombia

- Los datos son recolectados cada 15 días, mediante el API de mercadeo de Facebook. Los estimados mostrados estan basados en los filtros utilizados y en el comportamiento de los usuarios en los últimos 30 días. Obtén mas información sobre el alcance potencial
- Los estimados presentados no están diseñados para coincidir con censos u otras fuentes oficiales. Facebook no provee censos digitales o conteos de migrantes y/o refugiados. Estos estimados deben ser vistos como una señal par ser utilizada en triangulación.
- Facebook solo provee la definición del comportamiento seleccionado (Expatriados Venezuela). No provee datos estadísticos ni históricos
- El comportamiento depende de la información proporcionada por el usuario en Facebook, su ciudad actual y ciudad de origen y la estructura de la red de amigos (por ejemplo, tener al menos dos amigos de Facebook en el país de origen y dos amigos de Facebook en el país de destino). Leer más: Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants, ZAGHENI, Emilio. WEBER, Ingmar. GUMMADI, Krishna

Changes to Facebook's Backend



What About Ukraine?



Nowcasting Daily Population Displacement in Ukraine through Social Media Advertising Data
https://onlinelibrary.wiley.com/doi/full/10.1111/padr.12558

So, what is it good for?

- Monitor migration and population shifts
  Where are more (or less) users than usual?

- Map digital gender gaps
  Where do women have less internet access? Where are digital skill gaps?

- Trackings shifts in interests
  Who is (not) interested in climate change? Or in blockchain technologies?

Works not only for Facebook but for *_any_* platform that has targeted ads.

What could go wrong?

Work Reaches Beyond Academia … Is that Good?

Sage PolicyProfiles

Log in

Ingmar Weber

134

citations across 80 policy documents

47

policy documents which cite this researcher's work have been cited a further 602 times in 500 other policy documents

6

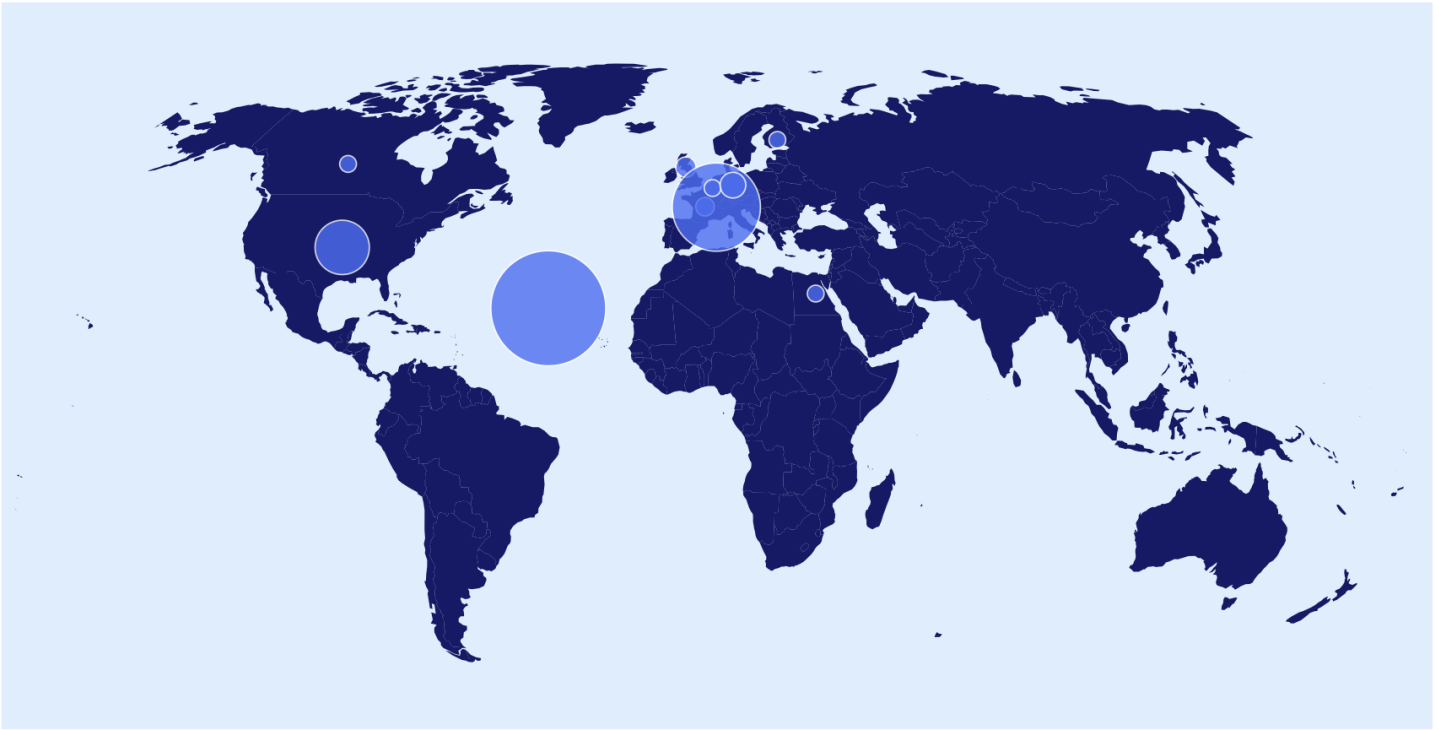
name mentions in 6 policy documents

What's the definition of a policy document?

Map | Timeline

Any time

Export map



https://policyprofiles.sagepub.com/profile/26375/ingmar-weber

The Risk of Counting People

- Hitler came to power in January 1933
- In April 1933 a census is announced
- Sub-goal: identify Jews, Roma, and other minorities
- Punch cards and IBM facilitated tabulation
- Outcome: 500k -> 2M Jews identified based on ancestry
- Better data helped to “optimize” the holocaust

So, what about the Rohingya? Can we map them?

Or Mexicans in the US?

Cui Bono – Fortress Europe?

FRONTEX

Lead Data Scientist AD8

Details of this vacancy

Reference No	RCT-2023-00049
Deadline (midday, 12:00 h of Warsaw local time)	21/08/2023
Documents	*Application form Lead Data Scientist AD8*
	Vacancy Notice Lead Data Scientist AD8 in IFC
Status	Recruitment closed

https://microsite.frontex.europa.eu/en/recruitments/RCT-2023-00049

https://www.opendemocracy.net/en/podcasts/podcast-borders-belonging/europe-migration-crisis-fortress-wall/

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

SIC Saarland Informatics Campus

Compute. Collaborate. Change.

An appeal for more interdisciplinary research

Data Needs Context

Computer scientists think in numbers, social scientists think in people

Wicked Problems Need New Approaches

[Insert what matters to you] will not be tackled by any single discipline

Responsible AI Needs More than Computer Science

Responsible AI development, deployment and regulation is only possible with informed exchanges

Talent Needs a Mission

An opportunity to attract non-traditional profiles to computer science and research

Key References

Computational Social Science from Space; IC2S2 2023 Tutorial, S. Šćepanović and I. Weber; https://sites.google.com/view/cssfromspace

Estimation of Internal Displacement in Ukraine from Satellite-Based Car Detections; Scientific Reports 2024, M.-C. Rufener, F. Ofli, M. Fatehkia, and I. Weber; https://www.nature.com/articles/s41598-024-80035-8

VME: A Satellite Imagery Dataset and Benchmark for Detecting Vehicles in the Middle East and Beyond; Scientific Data 2025, N. Al-Emadi, I. Weber, Y. Yang, and F. Ofli, https://www.nature.com/articles/s41597-025-04567-y

A Weak Supervision Learning Approach Towards an Equitable Mobility Estimation; ICWSM Workshop on Data for the Wellbeing for the Most Vulnerable 2025, T. Aidoo, T. Koebe, A. Maurya, H. Shrestha, I. Weber; https://workshop-proceedings.icwsm.org/abstract.php?id=2025_04

Monitoring of the Venezuelan Exodus through Facebook's Advertising Platform; PLOS ONE 2020, J. Palotti, N. Adler, A. Morales-Guzman, J. Villaveces, V. Sekara, M. G. Herranz, M. Al-Asad, I. Weber; https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0229175

Nowcasting Daily Population Displacement in Ukraine through Social Media Advertising Data; Population and Development Review 2023, D. R. Leasure, R. Kashyap, F. Rampazzo, C. Dooley, B. Elbers, M. Bondarenko, M. Verhagen, A. Frey, J. Yan, E. Akimova, M. Fatehkia, R. Trigwell, A. Tatem, I. Weber, M. Mills; https://onlinelibrary.wiley.com/doi/full/10.1111/padr.12558

More related work at https://ingmarweber.de/publications/

Thanks!

Happy to show demos of satellite providers or advertising platforms during the Q&A.