

Internet History Initiative

Overview and Exhortation

Jim Cowie

Berkman Klein Center for Internet & Society at Harvard University

IRTF / Global Access to the Internet for All

IETF 121, Dublin

6 Nov 2024

Gilbert Simondon (1924-1989)

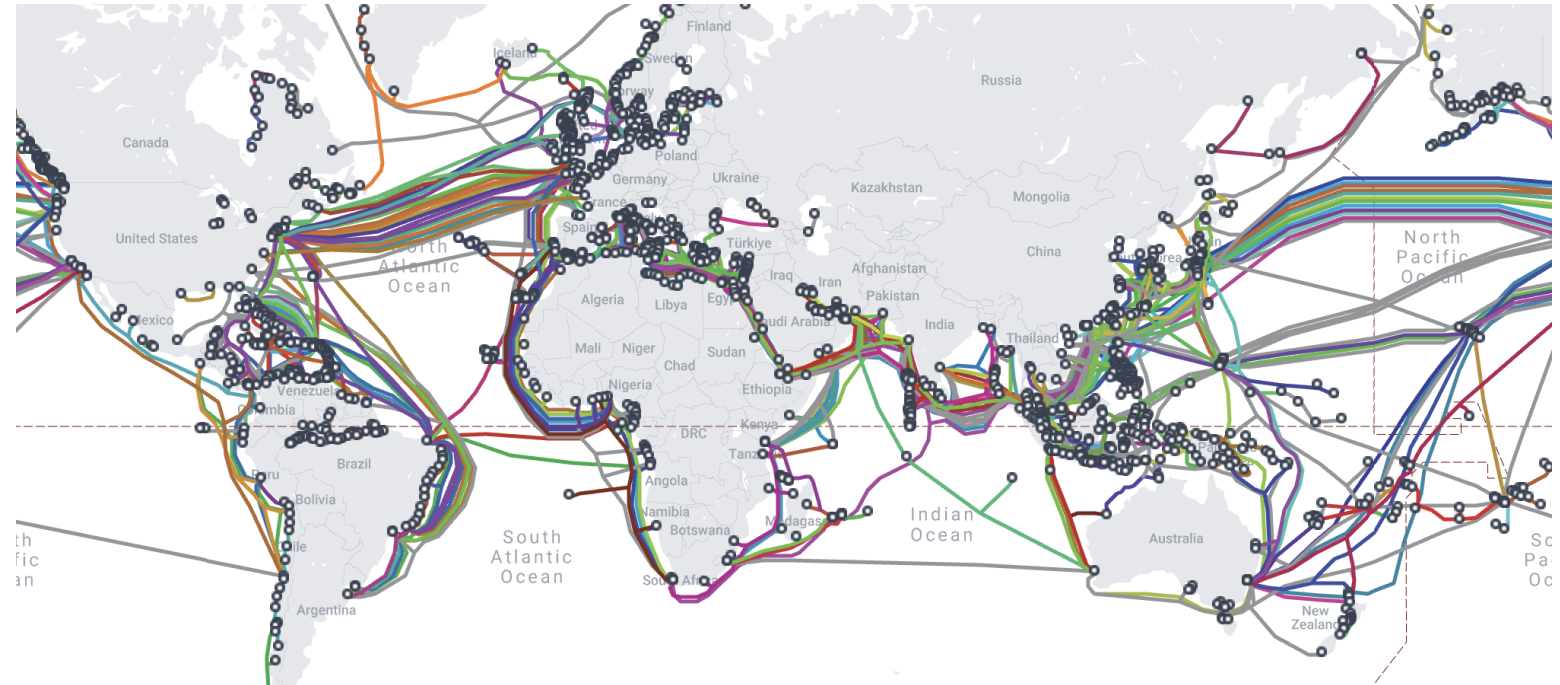
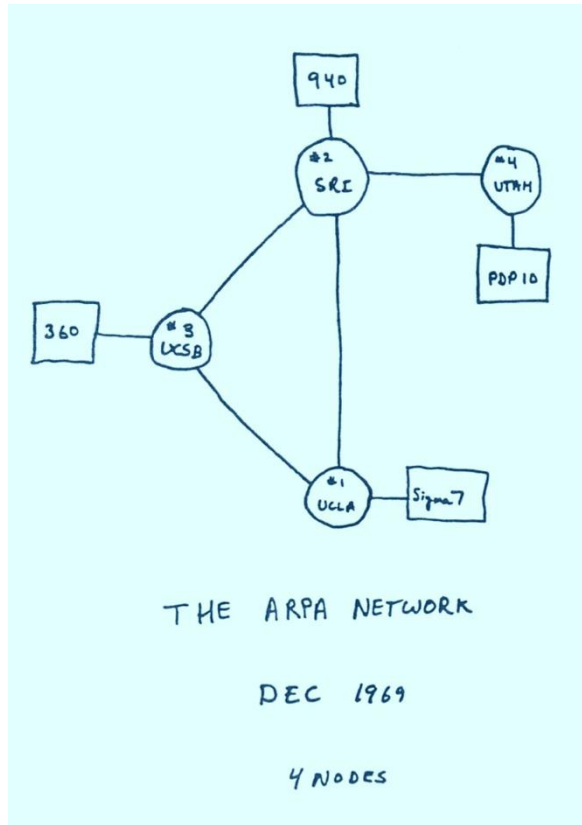
"Transforming all the conditions of human life, augmenting the exchange of causality between what man produces and what he is, true technical progress might be considered as implying human progress if it has a **network structure, whose mesh is human reality**"

-- "The Limits of Human Progress: A Critical Study" (1959)



https://en.wikipedia.org/wiki/File:Lemonde-Gilbert_Simondon.jpg

The Internet's Growth Took One Path of Many...



...and this may not be the best of all possible worlds.

Paths Not Taken

Internet Draft

Internet Architecture Board
July 1992

Expires: January 1993

IP Version 7
DRAFT 8

Abstract

Internet growth has created serious problems of address space consumption and routing information explosion. A solution to these problems requires a new version of the Internet Protocol, which we call IP version 7 ("IPv7"). This memo presents architectural guidelines that any IPv7 should meet. It then discusses how an IPv7 based upon the OSI CLNP protocol would meet these requirements, and presents the reasons for the IAB's preference for this solution. Finally, it makes a three-part recommendation: (1) proceed at full speed on CIDR; (2) do the design work on IPv7 based on CLNP; and (3) continue to pursue research in advanced routing and other future extensions of the Internet architecture.

End of the End-to-End Internet?

- “Success breeds success” models (or just multiplicative growth compounding **first-mover advantage**) created our current world
- A few (largely US-based) platform companies “up the stack” have now arrived one hop away from all the users they care to serve, worldwide, making **sophisticated routing irrelevant**
- These companies can afford **plenty of old-growth address space** in which to host everyone’s applications, especially now that TLS multiplexes many named services onto few IPs
- Communities that join today’s Internet don’t need ‘real’ addresses any more because we are all **merely consumers, not providers**
- Game over **(play again y/n?)**



Welcome to the Internet History Initiative

[Latest News](#) 

[Connect, Discuss](#) 

[Project Overview](#) 

Preservation, curation, and celebration of the Internet's historical datasets



Basic Motivation

- The Internet is a game that can be played over and over
 - *...If it remains open to end-to-end innovation*
- Threats to permissionless innovation come from many sources
 - *...The best defense is to build allies from many sources*
- The Internet community has not always been great at telling compelling, actionable stories about its own value+history
 - *....Let's work on that, together.*

A Basis for History, In Brief

- Engage institutional holders of the Internet's historical content and datasets (some at right; no endorsement implied!)
- Discuss potential constraints on preservation: engineering, scaling, resources, privacy, lack of demand by researchers
- Bring research librarians to the table as well.
- Identify solutions and create templates for action that data holders can appeal to when it's time to make hard decisions about getting rid of their old datasets
- **Preserve the data. Replicate the data. Curate the collection.**
- Engage scientists and storytellers who can use the collection to explain why the Internet way of networking should be a net benefit to all of us, and **in particular, to the newly connected**



Internet History: Many Lenses

- Oral culture
 - Ethnographies
 - Interviews
- Digital archives
 - IETF Datatracker
 - Project webpages
 - Listserv archives

- Measurement history
 - BGP, DNS, ICMP, ...

Internet History Initiative

It will take **real historians**, and the **passage of time**, to generate the frameworks within which to assemble competing interpretations of the available evidence, to figure out what happened to humanity over the last 50 years, and what role the Internet played in those outcomes.

(Some Of) The Data

IXP/Peering
(PeeringDB Archives)

Page Source IPs
(Internet Archive)

DNS Names
(CAIDA 2008-)

Root Zones, DITL
(DNS-OARC)

RDNS
(RIPE 2009-)

DNS Resolutions
(RIPE Atlas 2012-)

PingER
(SLAC 1997-2024)

Active Measurement
(CAIDA Ark 2015-)

Performance
Measurements
(MLAB 2009-)

Ping/Trace
(RIPE Atlas 2012-)

BGP Routing
(Routeviews 1999-)

BGP Routing
(PCH 1997-)

BGP Routing
(RIPE RIS 1999-)

Regional Registry Allocations (ASNs, IPv4, IPv6 address space)

Example: What We Stand To Lose

- PingER project collected continuous data from 1997 to May 2024
- Estimates of latency, jitter, and bandwidth
- Specific attention to Global South universities and developing economies
- **Status: Extinct**
Endangered



Quantifying the Worldwide Digital Divide: The Emergence of Africa

Prepared by: R. Les Cottrell
SLAC, Stanford, USA
ISPA/iWeek, Pretoria, S. Africa, Sept 21, 2011
cottrell@slac.stanford.edu



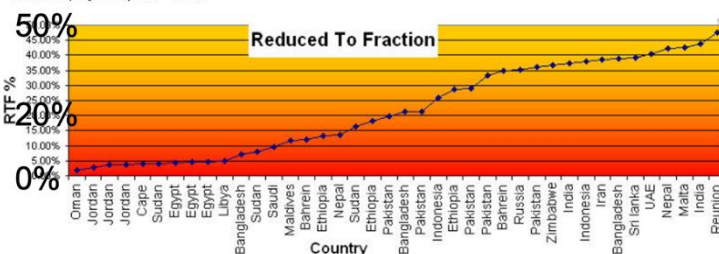
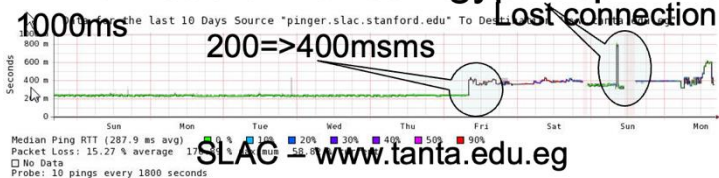
SEACOM is a 1.28Tbps, 15,000 kilometer-long undersea fiber-optic cable system that will provide high-speed communication between East Africa, South Asia and

<http://www.slac.stanford.edu/grp/scs/net/talk11/africa-sep11.pptx>

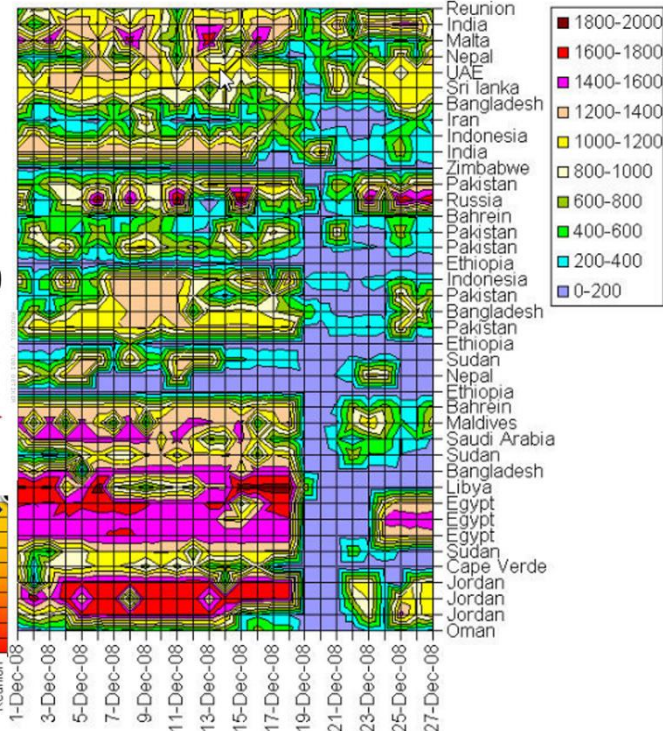
Early, Important Work

Example: Multiple routes important

- Not only for competition
- Need redundancy
- Mediterranean Fibre cuts
 - Jan 2008 and Dec 2008
 - Reduced bandwidth by over 50% to over 20 countries
- New cable France-Egypt Sep 1 '10



Derived TCP Throughput in kbits/sec from SLAC to Hosts in Countries Affected by the Mediterranean Fibre Cable Cuts December 2008



- Positive correlation between PingER throughput & IDI, especially for populous countries

- PingER measurements automatic

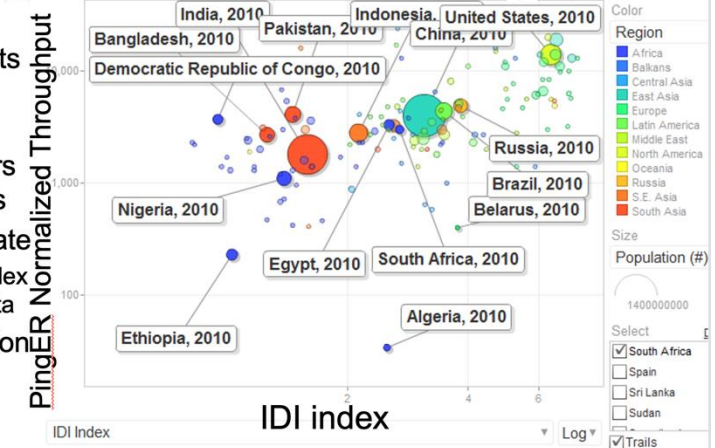
No army of data gatherers & statisticians

More up to date

- IDI 2009 index for 2007 data

Good validation

Anomalies interesting

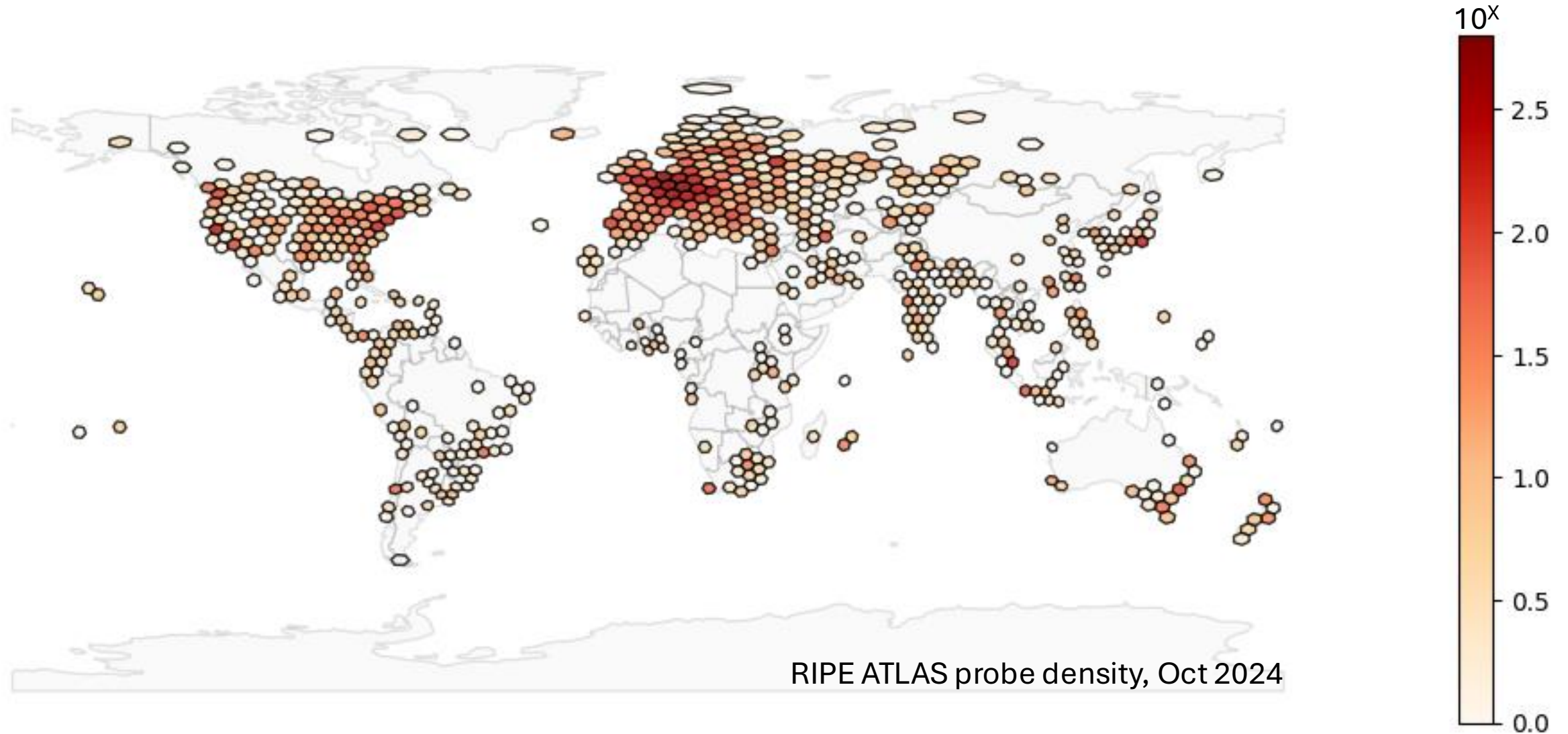


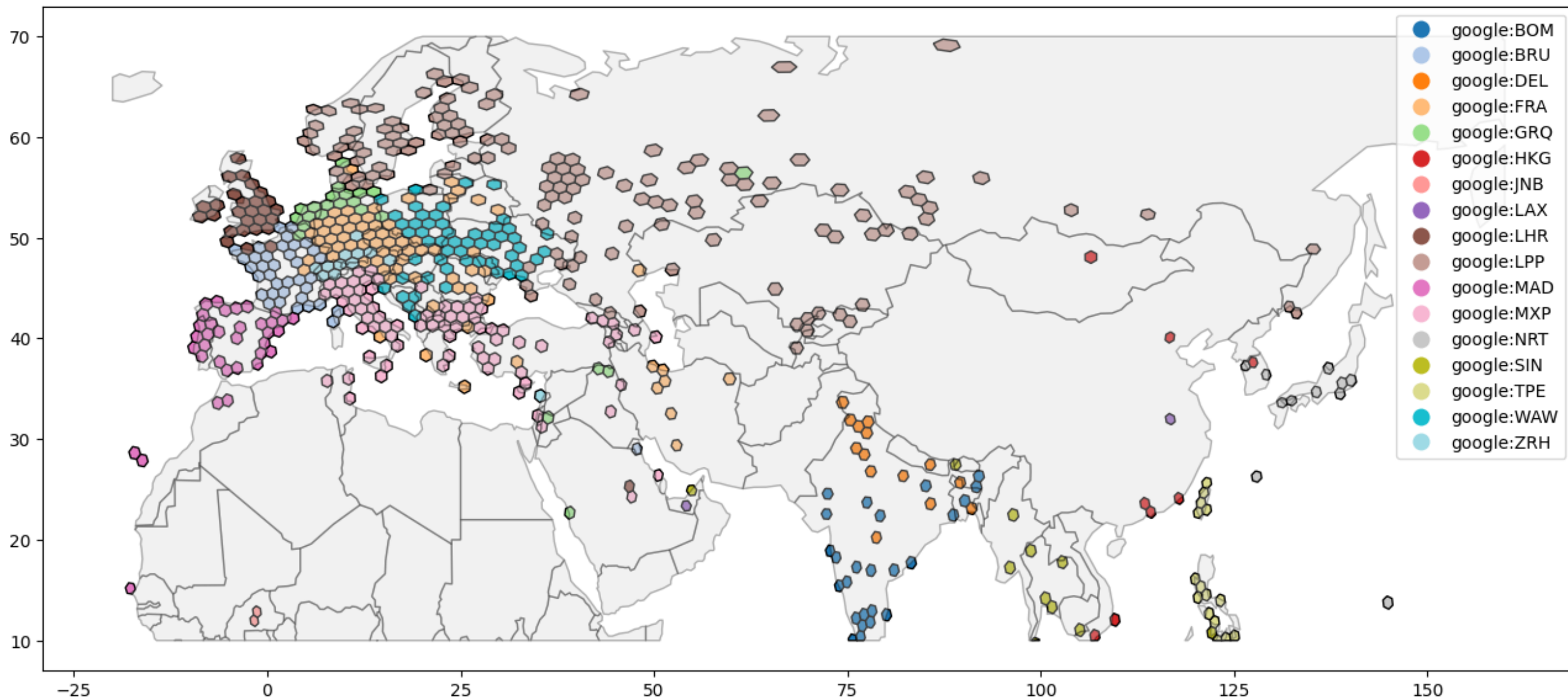
A copy of this data was revived by SLAC IT in September 2024

It's being mirrored to IHI's S3 for intermediate preservation

Next steps: curation, transcription to modern data formats

Network Measurement: Need More Global South





“Watersheds” for Google Public DNS (8.8.8.8)

Sparse resolution due to probeset bias; can see some interesting long hauls

Engaging Social Scientists

American Economic Review 2019, 109(3): 1032–1079
<https://doi.org/10.1257/aer.20161385>

The Arrival of Fast Internet and Employment in Africa[†]

By JONAS HJORT AND JONAS POULSEN*

To show how fast Internet affects employment in Africa, we exploit the gradual arrival of submarine Internet cables on the coast and maps of the terrestrial cable network. Robust difference-in-differences estimates from 3 datasets, covering 12 countries, show large positive effects on employment rates—also for less educated worker groups—with little or no job displacement across space. The sample-wide impact is driven by increased employment in higher-skill occupations, but less-educated workers' employment gain less so. Firm-level data available for some countries indicate that increased firm entry, productivity, and exporting contribute to higher net job creation. Average incomes rise. (JEL F14, J23, J24, J63, L86, O15, O33)

The Employment Effects of Mobile Internet in Developing Countries*

Gaurav Chiplunkar
University of Virginia

Pinelopi Koujianou Goldberg
Yale University, BREAD, CEPR and NBER

November 2022

Abstract

We examine the employment effects of 3G mobile internet expansion in developing countries. We find that 3G significantly increases the labor force participation rate of women and the employment rates of both men and women. Our results suggest that 3G affects the type of jobs and there is a distinct gender dimension to these effects. Men transition away from unpaid agricultural work into operating small agricultural enterprises, while women take more unpaid jobs, especially in agriculture, and operate more small businesses in all sectors. Both men and women are more likely to work in wage jobs in the service sector.

If Technology Has Arrived Everywhere, Why Has Income Diverged?

Diego Comin

Martí Mestieri*

July 20, 2017

Abstract

We study the cross-country evolution of technology diffusion over the last two centuries. We document that adoption lags between poor and rich countries have converged, while the intensity of use of adopted technologies of poor countries relative to rich countries has diverged. The evolution of aggregate productivity implied by these trends in technology diffusion resembles the actual evolution of the world income distribution in the last two centuries. Cross-country differences in adoption lags account for a significant part of the cross-country income divergence in the nineteenth century. The divergence in intensity of use accounts for the divergence during the twentieth century.

Keywords: Technology Diffusion, Transitional Dynamics, Great Divergence.

JEL Classification: E13, O14, O33, O41.

Looking outside the Internet community, we may be more likely to find career scientists with the perspective and distance to take advantage of the signals in our data history.

Conclusion: How You Can Help

- Open difficult conversations about preserving data
- This feels a lot like estate planning – nobody likes to think about it
- Pass along **leads for endangered datasets** before they are gone forever.
- Pass along **links to great papers** by evidence-based social scientists who might break some ground if given these datasets.
- Pass along **stories of local communities** that remember the arrival of the Internet, so that we can align those histories with the corresponding data histories as guideposts for future historians

Thank you!

<https://www.internethistoryinitiative.org>

Mastodon: @IHI@cooperate.social

Email: jacowie@cyber.harvard.edu