LEVERAGING BIG DATA AND "HUMAN AI" FOR INNOVATION AND DEVELOPMENT IN LATIN AMERICA AND THE CARIBBEAN

Emmanuel Letouzé, PhD
Director and co-Founder, Data-Pop Alliance
Marie Curie Fellow, Universitat Pompeu Fabra Barcelona
Fellow, MIT Connection Science, Visiting Scientist, Harvard Humanitarian Initiative
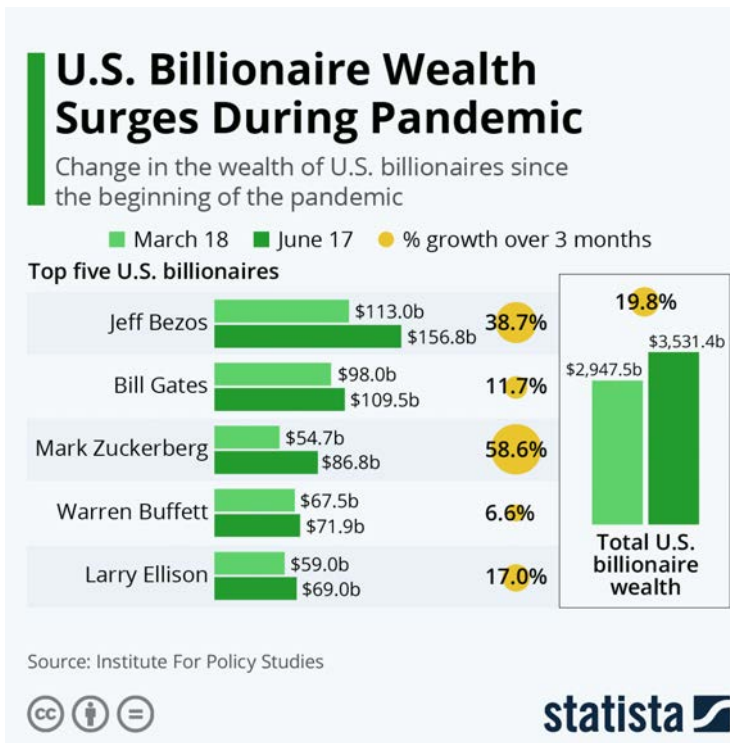Co-Founder, OPAL Project

September 20, 2021

Programme "Blockchain and Data Analytics for digital industries"

COVID exposed and exacerbated longstanding fault lines. As we enter the 2nd decade of the "4th IR", with rising inequalities, anxiety, distrust, fears, *will Big Data and AI help or hurt*?

Demonstrations in Latin America *in 2019*
Recent violent protests

HONDURAS
October 18
Opposition blocked streets in several regions to demand the departure of president Juan Orlando Hernandez

HAITI
Protest movement since end of August demanding resignation of president Jovenel Moise

PORT-AU-PRINCE
HONDURAS

ECUADOR
October 2-13
Protests over fuel price increases
8 dead

QUITO

BOLIVIA
October 22
Opposition claims "fraud" after presidential vote on October 20 and takes to streets to demonstrate

LIMA
LA PAZ

PERU
October 7
Hundreds of Peruvians demonstrate against the dissolution of parliament

BUENOS AIRES

SANTIAGO

ARGENTINA
September 24
General strike over economic policies of president Mauricio Macri

CHILE
October 18-23
Rise in train ticket prices sparks a wave of protests over inequalities
15 dead as of October 23

Sources: AFP bureaus
© AFP

THE WORLD'S RICHEST 1% HAVE MORE THAN TWICE AS MUCH WEALTH AS **6.9 BILLION PEOPLE.**

Source: Oxfam 2020

U.S. Billionaire Wealth Surges During Pandemic
Change in the wealth of U.S. billionaires since the beginning of the pandemic

■ March 18 ■ June 17 ● % growth over 3 months

Top five U.S. billionaires

Jeff Bezos	$113.0b / $156.8b	38.7%
Bill Gates	$98.0b / $109.5b	11.7%
Mark Zuckerberg	$54.7b / $86.8b	58.6%
Warren Buffett	$67.5b / $71.9b	6.6%
Larry Ellison	$59.0b / $69.0b	17.0%

Total U.S. billionaire wealth: $2,947.5b → $3,531.4b 19.8%

Source: Institute For Policy Studies

statista

IS DEMOCRACY DYING?

A WARNING FROM EUROPE: The Worst Is Yet to Come *by* ANNE APPLEBAUM

HOW AI COULD GIVE RISE TO TYRANNY *by* Yuval Noah Harari

Human Development Report 2019
UNDP
Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century

Location Data Says It All: Staying at Home During Coronavirus Is a Luxury

By Jennifer Valentino-DeVries, Denise Lu and Gabriel J.X. Dance April 3, 2020

#GEORGEFLOYD
#ICANTBREATHE
#SAYHISNAME
CAN'T BREATHE

Vodafone Institute
for Society
and Communications

DATA-POP
ALLIANCE

Using Data to Fight COVID-19

And Build Back Better

Emmanuel Letouzé
Nuria Oliver

Maria Antonia Bravo
Natalie Shoup

October 2020
Paper Series No. 2
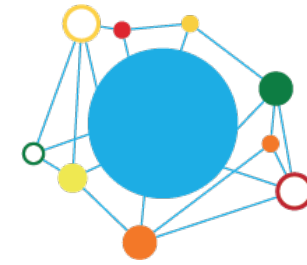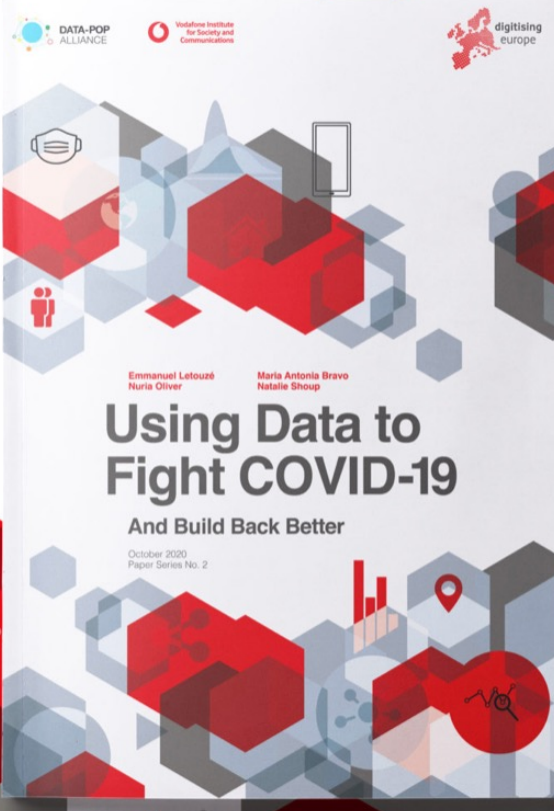
Social media companies and social platforms have a duty to the public to provide safeguards from theories that weaken trust in their governments and in science.

To realise this vision, four elements appear to be key:

One is context: we need to have a thorough understanding of the goals, implications and the impact on citizens and society of decisions in the longer term (from a science/tecUology, economic/commercial, social, political, legal and ethical point of view). It is also important to understand the different tecUologies being designed and used for real response, as well as the parameters and risks, benefits, limitations and impact of each. Furthermore, it is crucial to be mindful of the fact that not all responses can or must be digital, and that not all people will be able to access digital solutions. This means that solutions have to be thought in a holistic way so that everyone is included.

Another is education: citizens should be provided with clear, precise, under-

"The COVID-19 pandemic offers us a tremendous opportunity to leverage technology and data for positive social change. An opportunity that we cannot afford to miss. The time is now."

Emmanuel Letouzé, PhD
Director and Co-Founder
Data-Pop Alliance

[How] Can Big Data and AI both monitoring *and* promoting the SDGs and human development?

DATA-POP ALLIANCE
WORKING NOTE

Reflections on Big Data & the Sustainable Development Goals:
Measuring & Achieving Development Progress in the Big Data Era

INPUT TO THE BIG DATA AND SDGS CHAPTER OF THE 2015 GLOBAL SUSTAINABLE DEVELOPMENT REPORT

February 2015

"Big Data and Artificial Intelligence for Measurement and Evaluation of Human Development"

Emmanuel Letouzé, PhD
With contributions from Julián Riveros Clavijo, Maria Antonia Bravo and Magdalena Arbeláez Tobón

Background Paper to the 2019 Human Development Report on Inequalities

DATA-POP ALLIANCE

Human Development Report 2019

Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century

UNDP

2010-2015 and since : From big data to Big Data
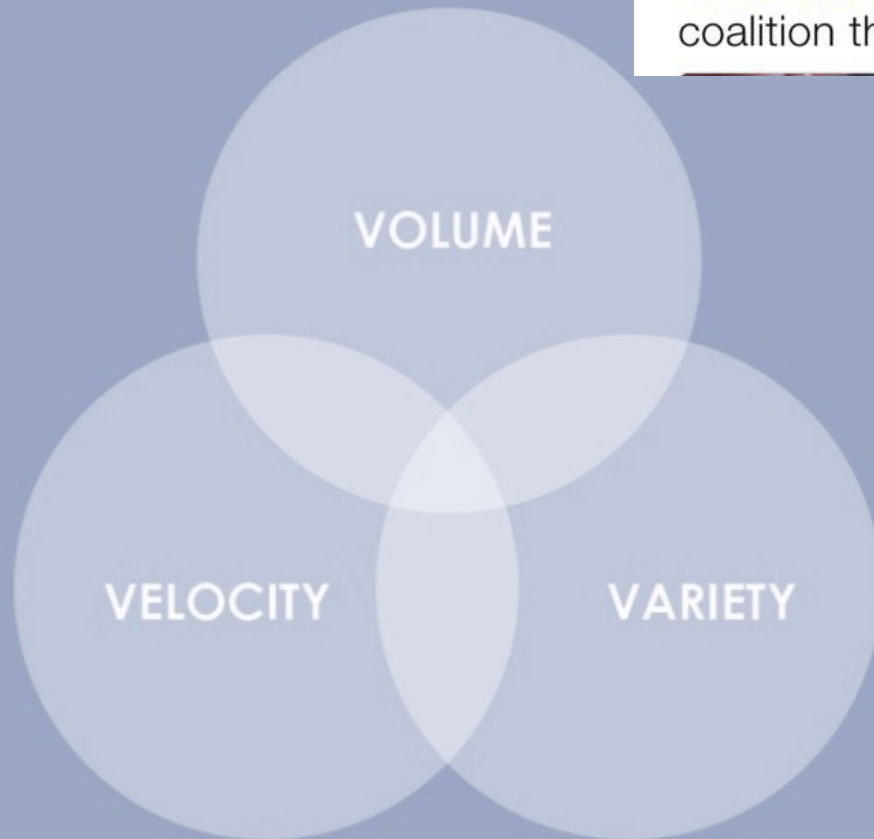
What would it look like and take to use 'AI' as both an *inspiration* and an *instrument* to build better future societies? Sandy Pentland's vision of a Human AI

TOWARDS A HUMAN ARTICIAL INTELLIGENCE FOR HUMAN DEVELOPMENT

Emmanuel Letouzé[1], Alex Pentland[2]

[1]Data-Pop Alliance, MIT Media Lab, and OPAL, [2]MIT and Data-Pop Alliance, and OPAL

Abstract – This paper discusses the possibility of applying the key principles and tools of current artificial intelligence (AI) to design future human systems in ways that could make them more efficient, fair, responsive, and inclusive.

http://datapopalliance.org/wp-content/uploads/2019/02/HumanAIITU2018-15.pdf

← Today this looks crazy to us!

Tomorrow this will look crazy to our kids? →

With AI—new data and tools—machines are learning. Could we learn from and with them?
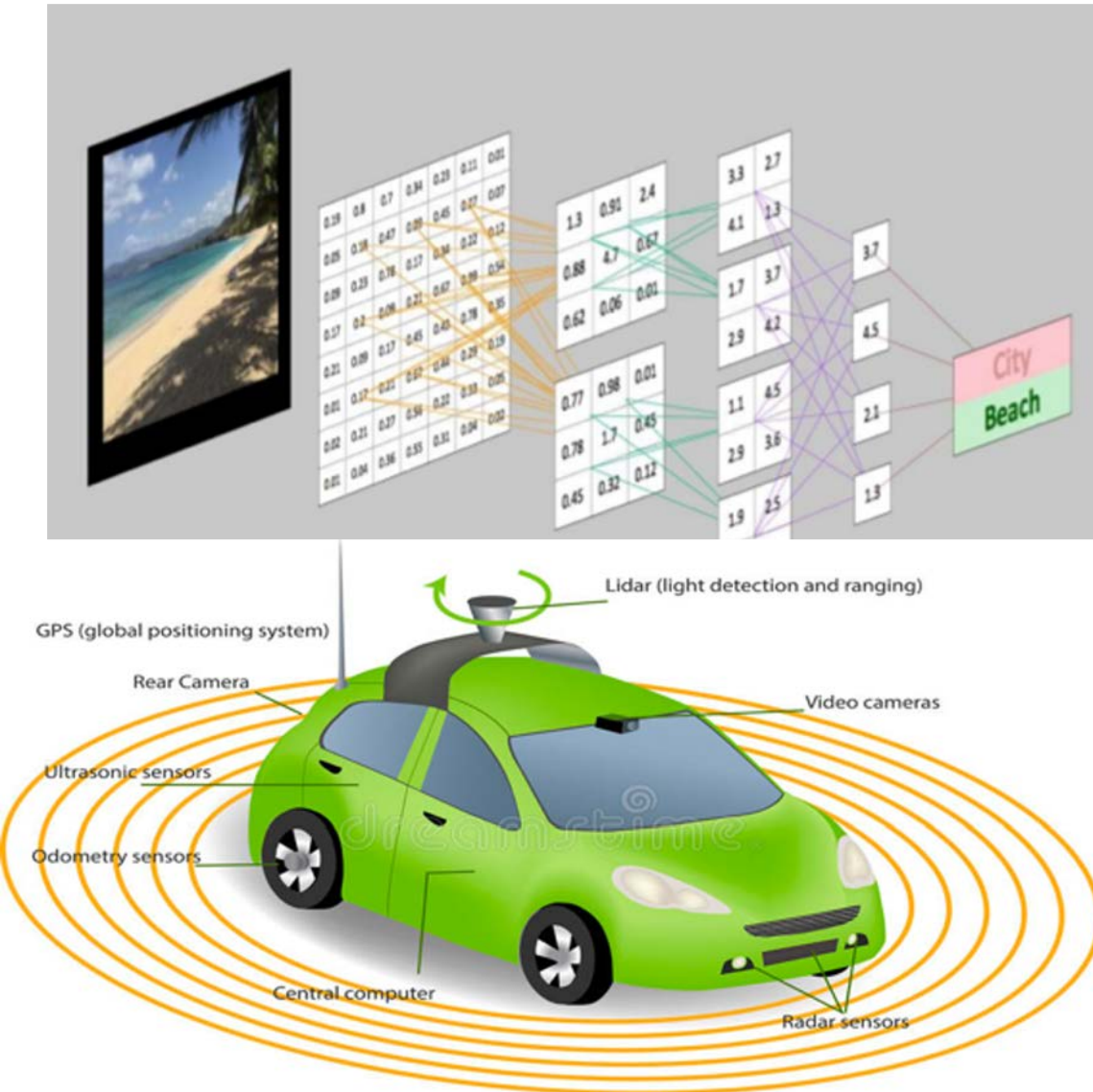
1. AI is at least **60+ years old**. **It is still pretty stupid.** We are **far from general AI**
2. The key novelty is of course **new DATA! (+ new tools; deep learning).**
3. The "good magic" of AI magic is the **credit assignment function=="What determines success vs failure?" " What should I do more of?"**
4. And **humans must remain in control; "in the loop"**

Human development, progress, innovation… do not happen linearly but via a series of "learning loops"
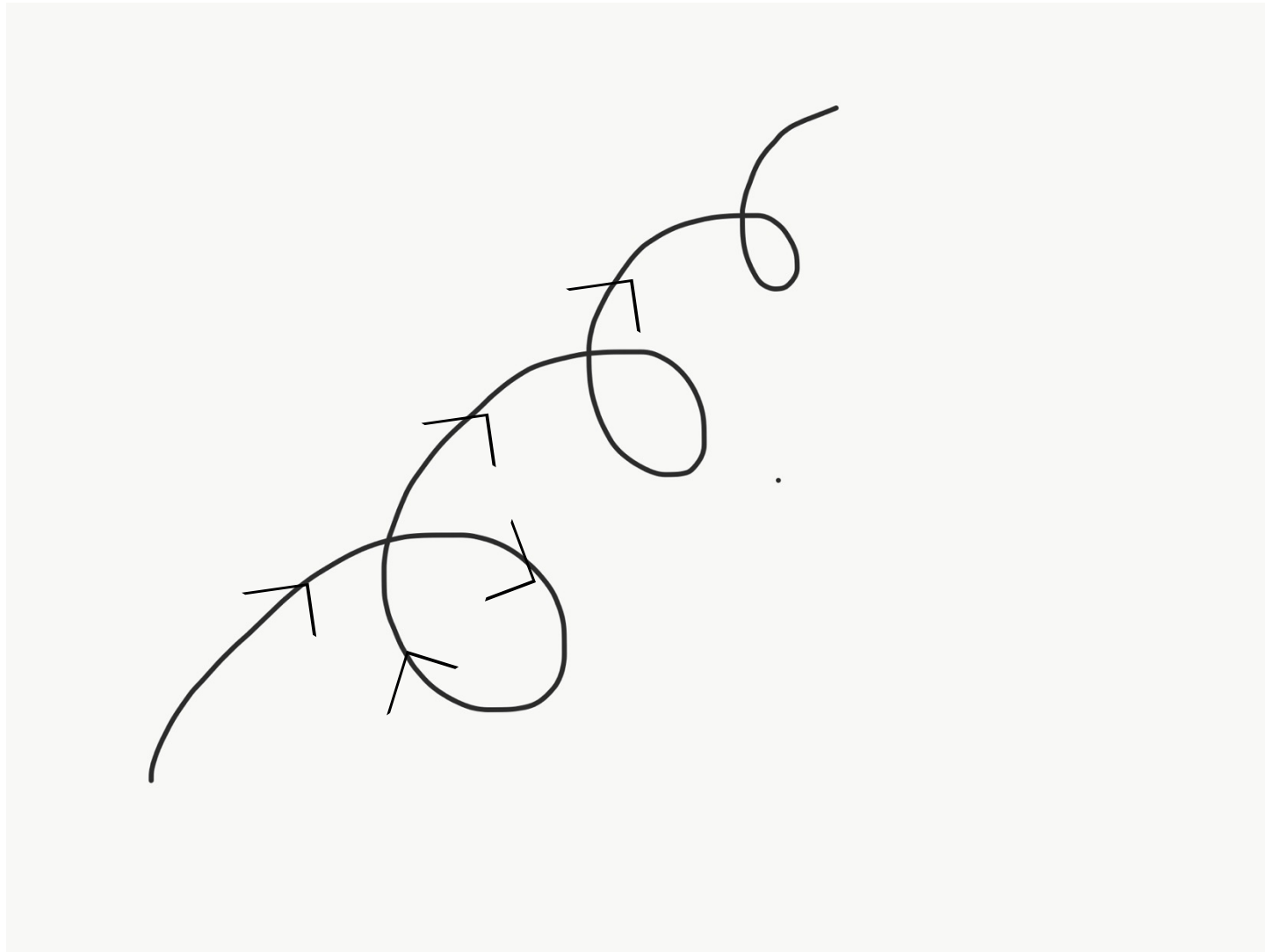
This process can be drawn and thought about as a function of confidence

Performance

Zoom on 1 loop and
its 4 zones / modes
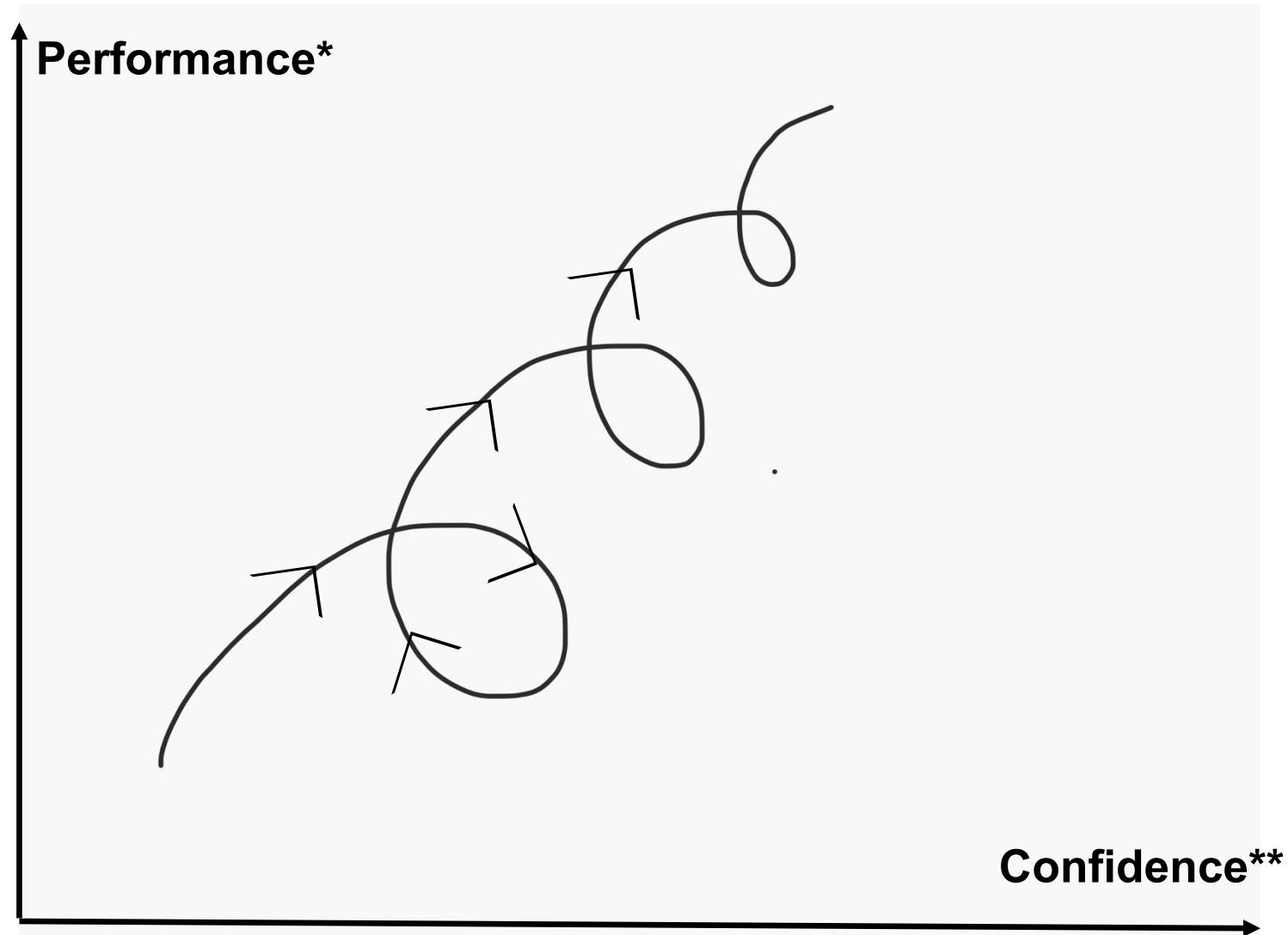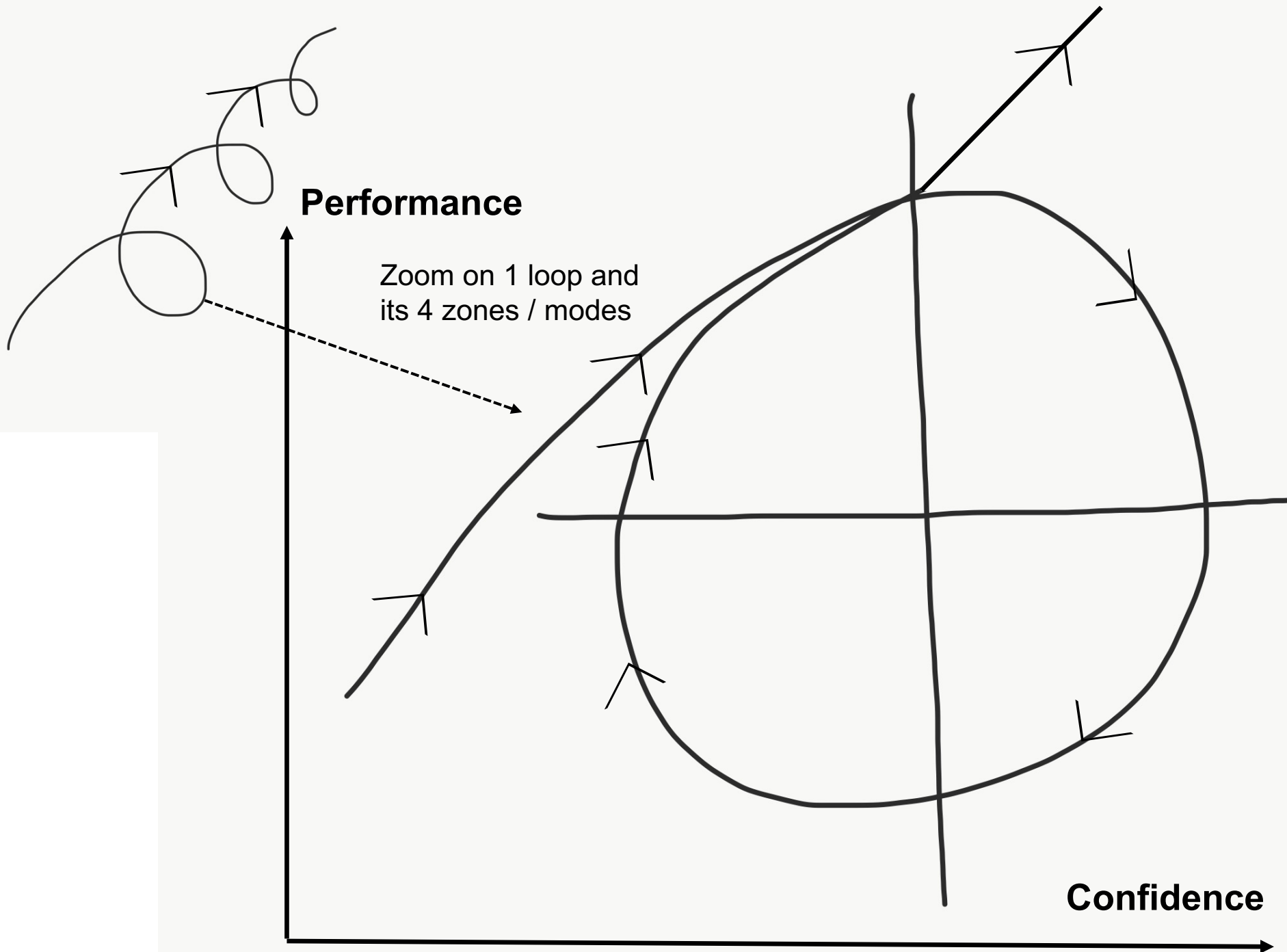
Confidence

Performance

Perf and Conf are **positively correlated** here. Confidence goes up and Perf goes up too. *Consolidation mode.*

Confidence

Performance

Confidence

Perf and Conf are **negatively correlated** here. More conf goes hand in hands with decreasing perf. *Overconfidence (arrogance?) zone / mode.*
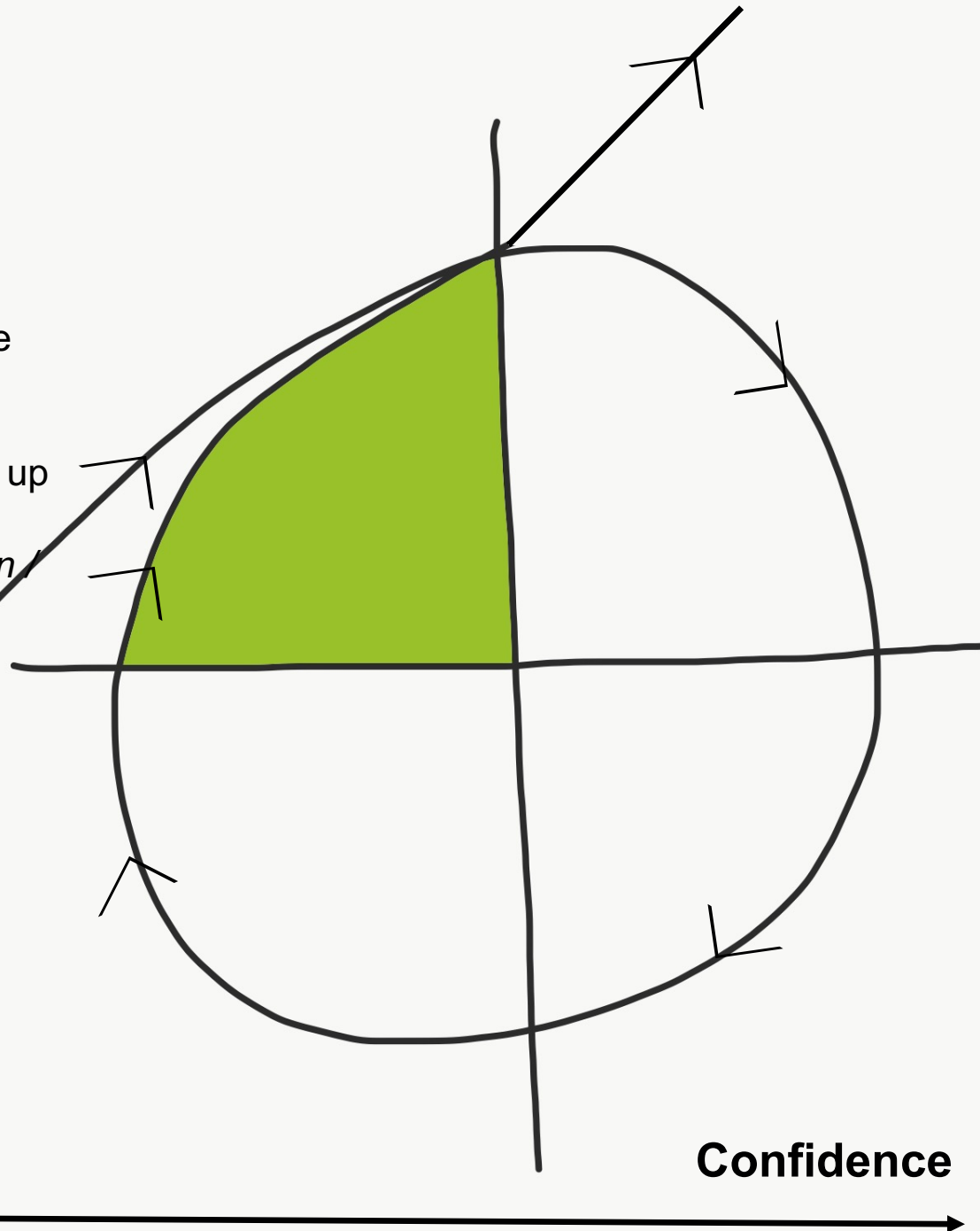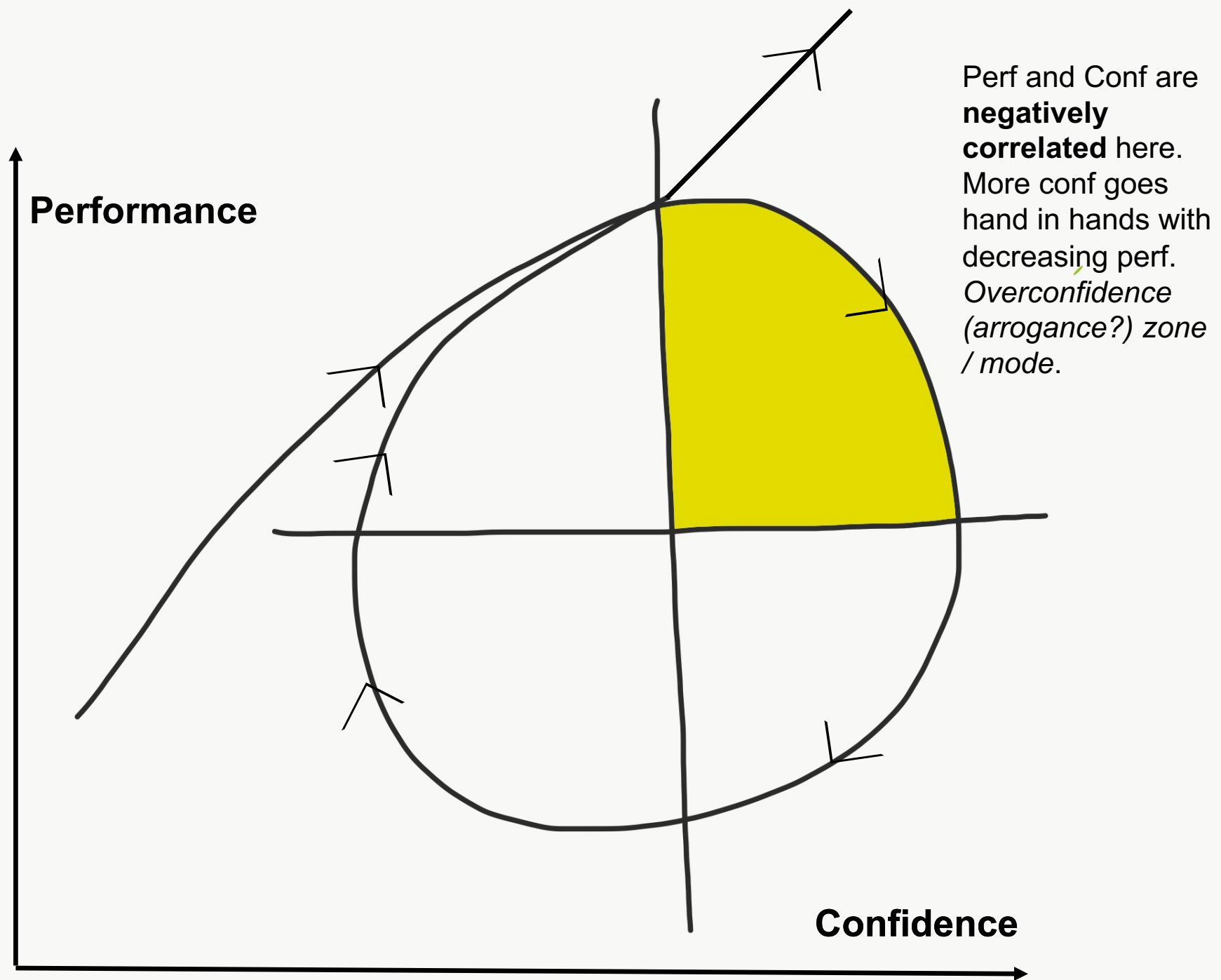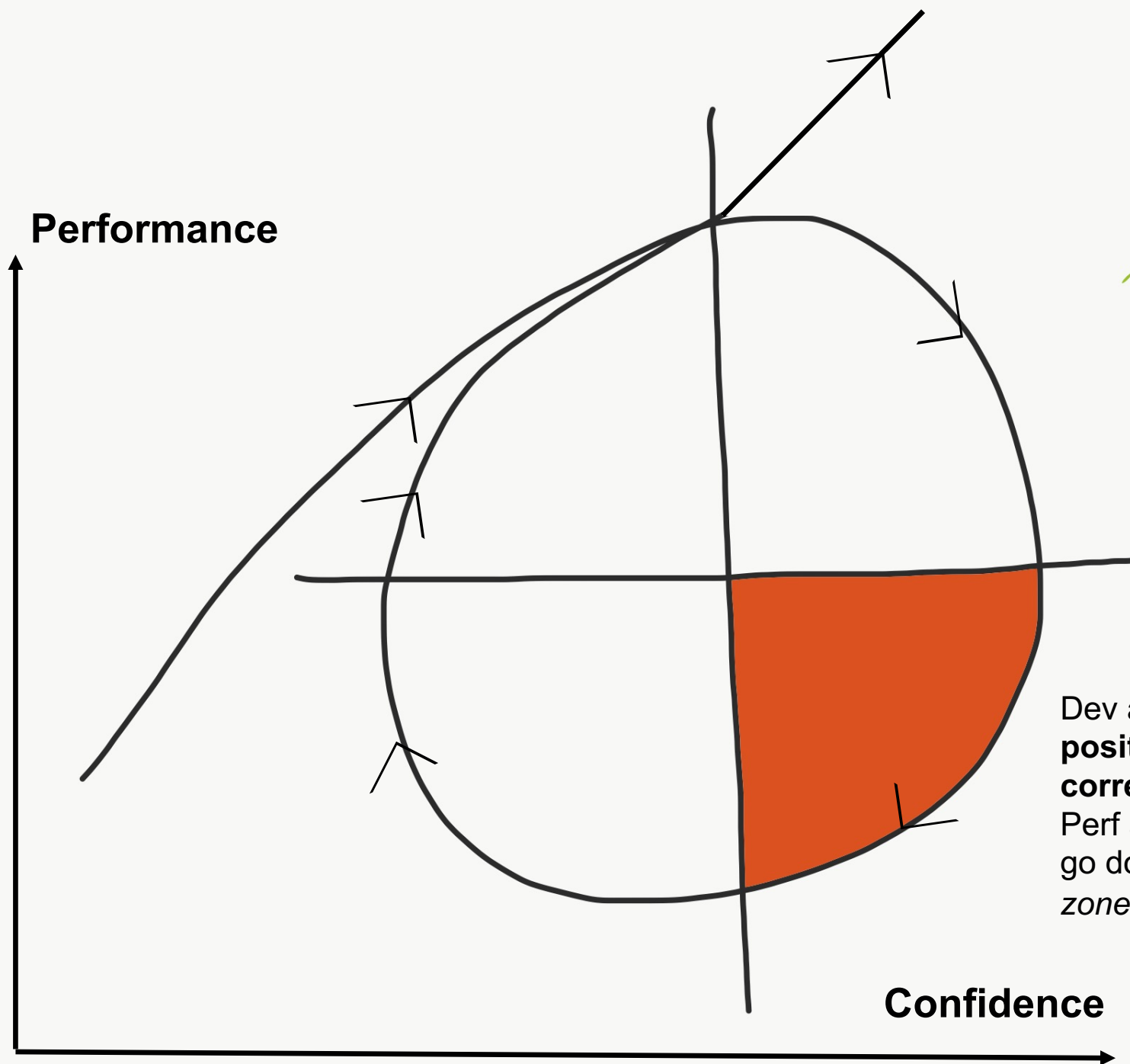
Performance

Dev and Perf are **positively correlated** here. Perf and dev both go down. *Crisis zone / mode.*

Confidence

Development

Confidence

Perf and Conf are **negatively correlated here**. Conf keeps going down but Perf goes back up. New systems happen. *Redemption.*
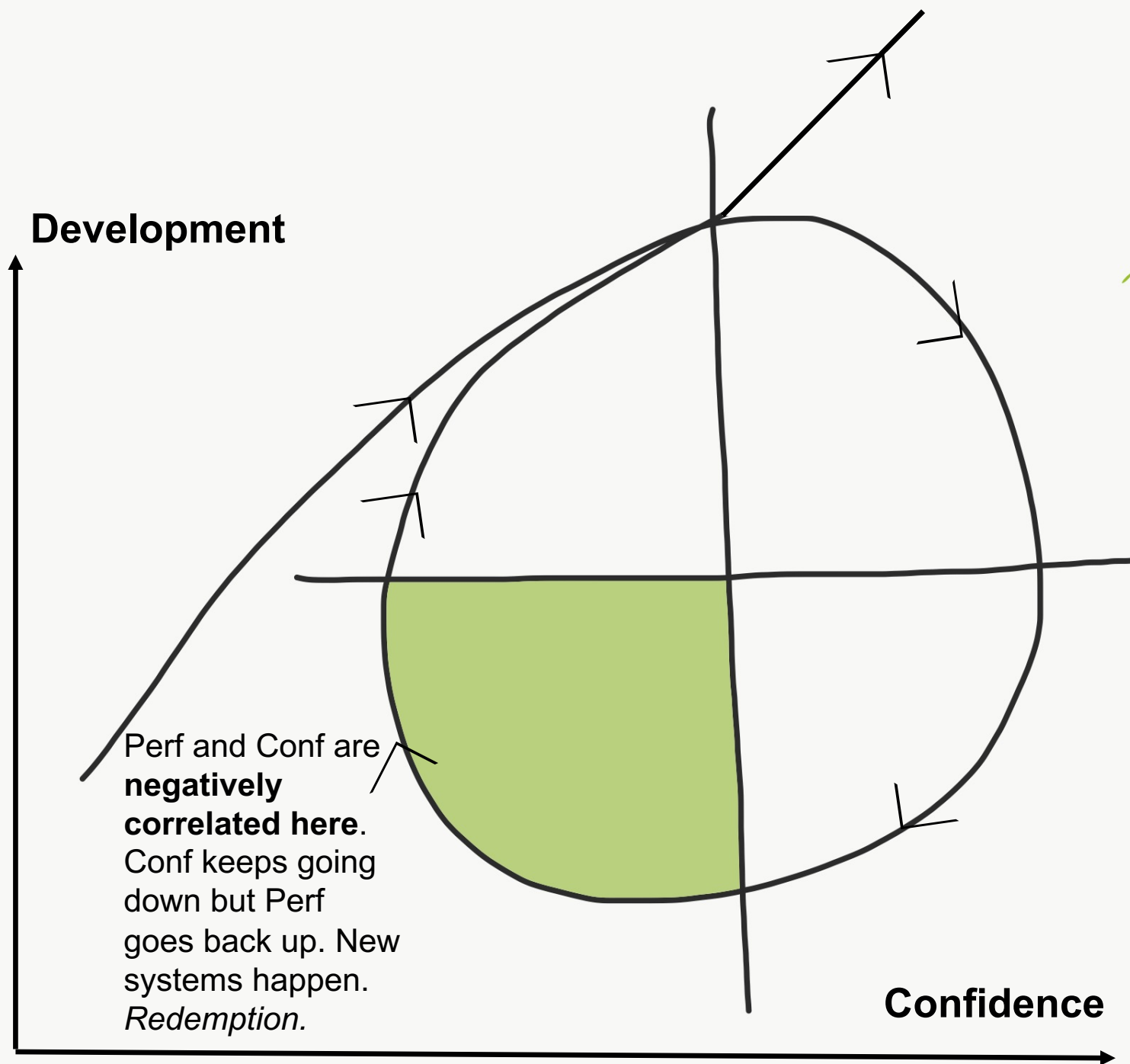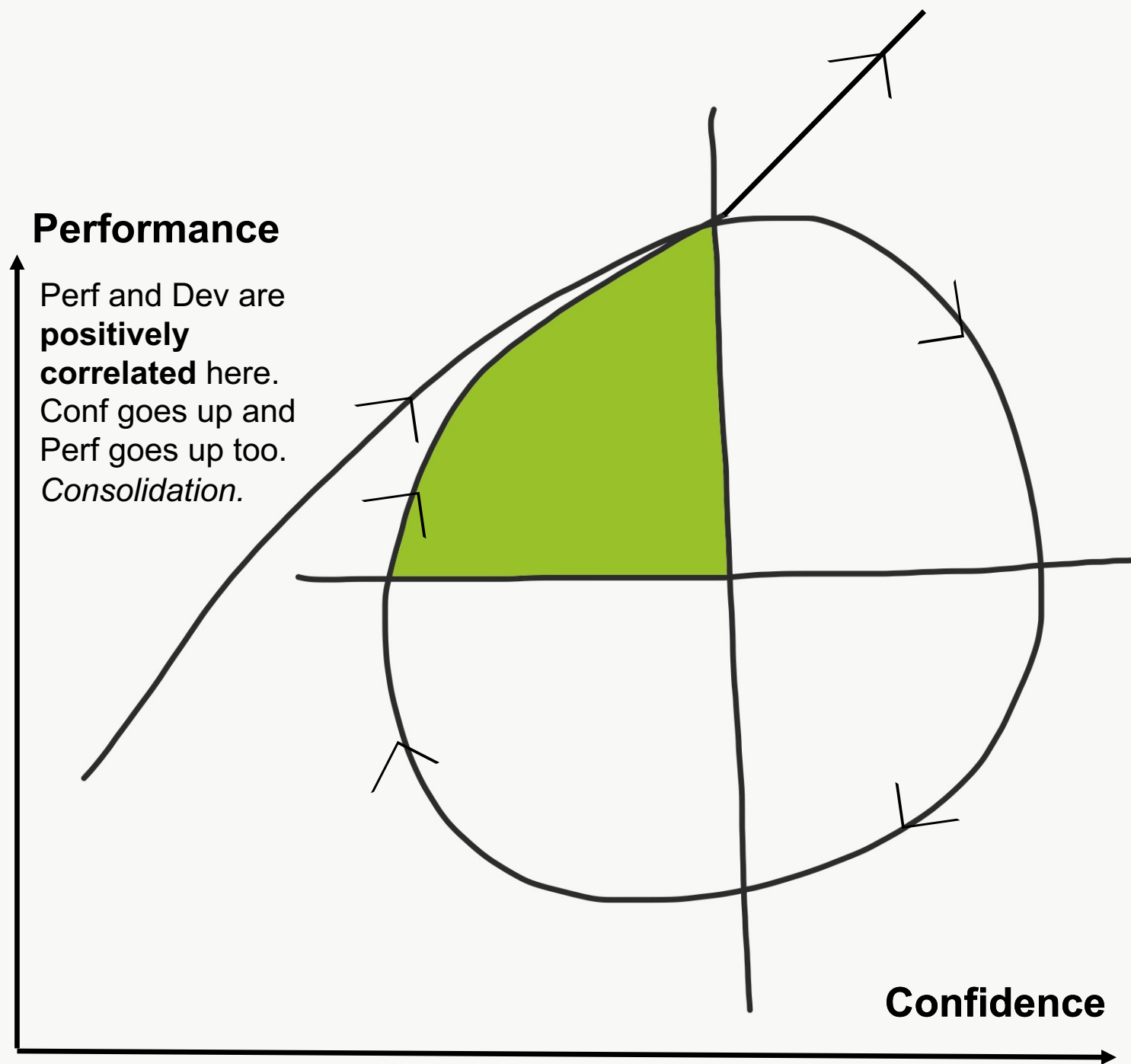
Performance

Perf and Dev are **positively correlated** here. Conf goes up and Perf goes up too. *Consolidation.*

Confidence

A few more observations about this system…
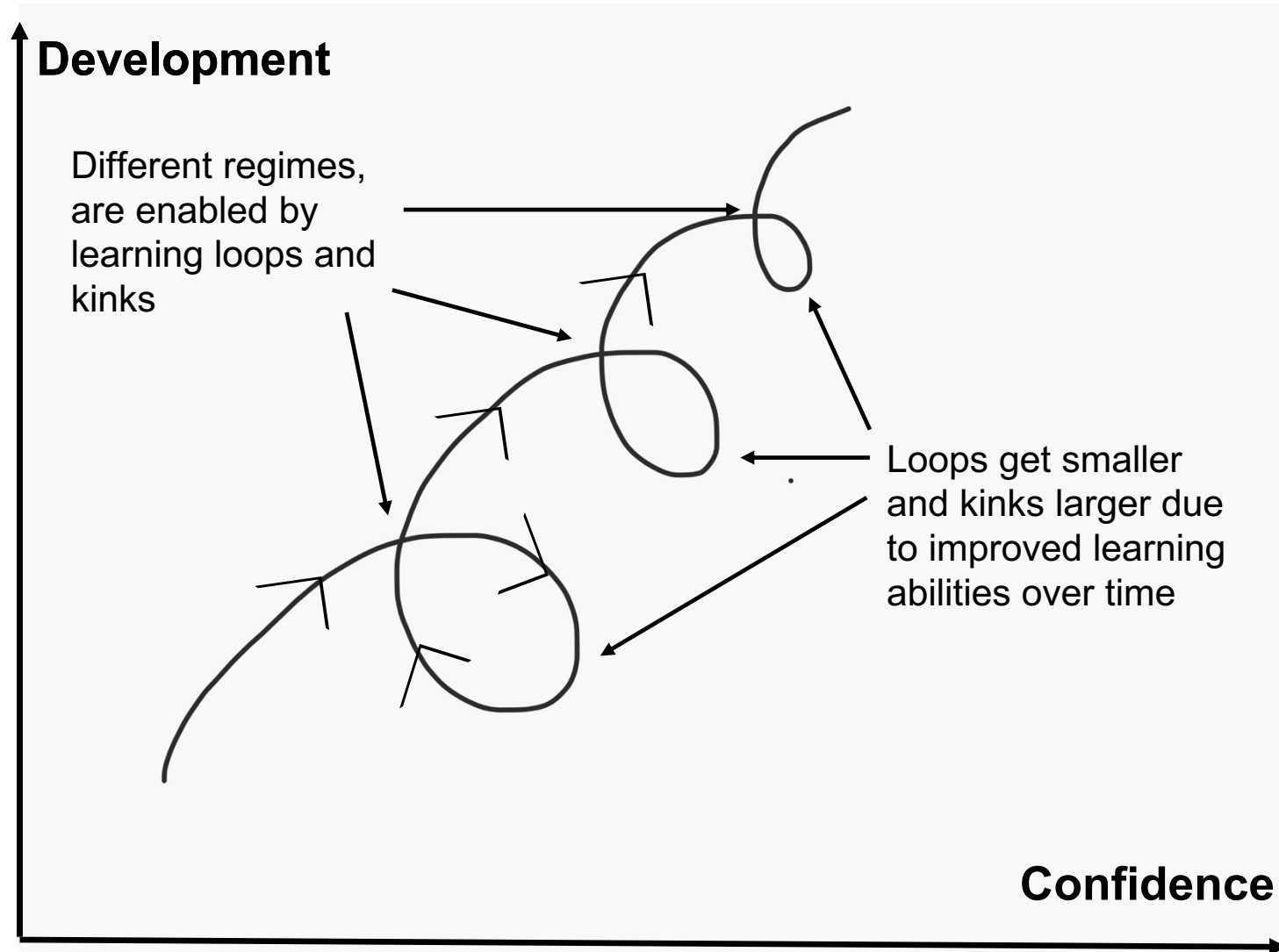
Main hurdles, risks and requirements on the path to a "Human AI for Human Development"

1. Political: weak incentives **economic and political elites** who benefit from status quo (even or especially as they want to be "do-gooders" :)
2. Sociological / psychological: distrust, disdain, echo chambers, alternative facts, that **hamper cooperation, consensus, compromise when it is already so hard to change people's minds and habits**…
3. Scientific / technical: lack of systemic and safe **data access**, connections, capacities…
4. Ethical / legal: can become "**Orwellian", tech-and-elite centric, intrusive**, all-too-prescriptive, self-perpetuating…
5. Commercial: where is the $$$$?
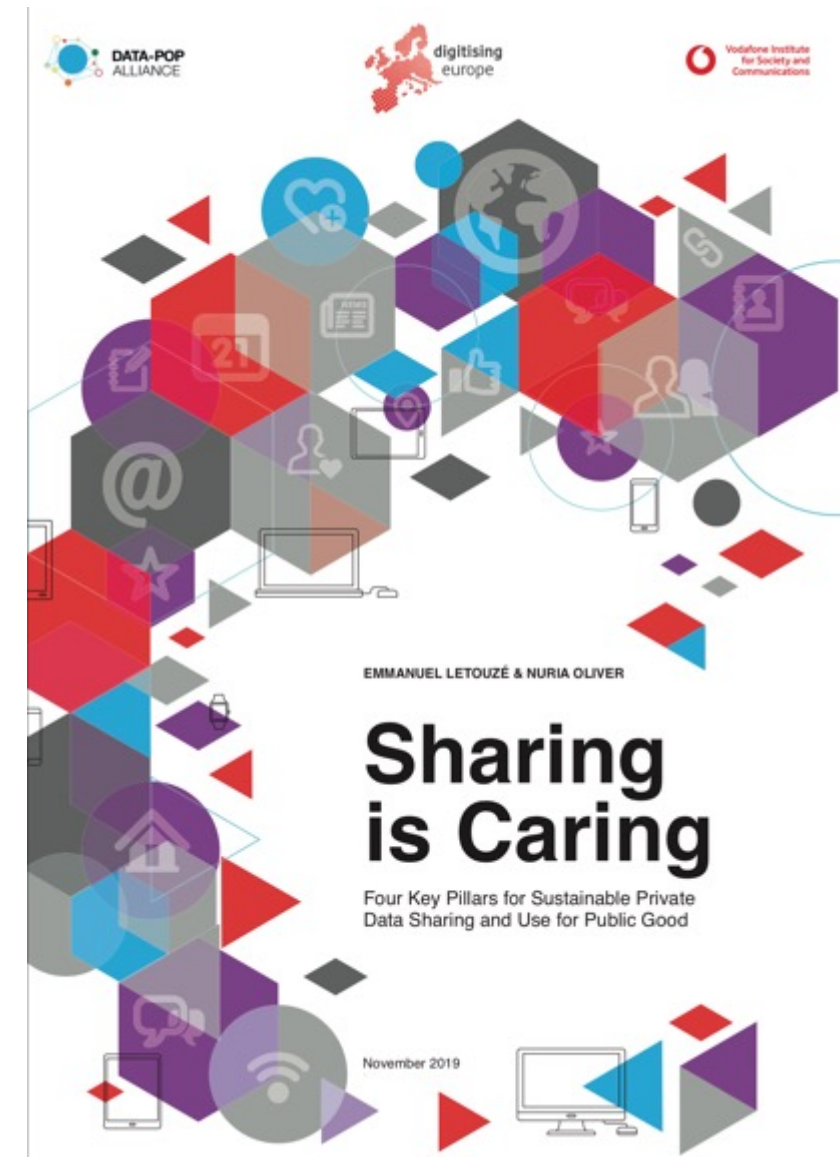
Many countries have started developing and implementing various national strategies, including in LAC

- ● E-Government / Digital Government
- ● Big Data / AI
- ● Ethic / Regulations

UE - Plan de Acción Europeo de Administración Electrónica 2011-2015 ●

Rep. de Corea - Portal de datos abiertos data.gov.kr ●

México - Estrategia Digital Nacional (EDN) ●

UE - Portal de datos abiertos de la UE ●

Estudio hacia una estrategia de IA en México ●

"Un plan de desarrollo de la inteligencia artificial de próxima generación" ●

Estrategia pancanadiense de IA ●

Estrategia Nacional de IA ●

Política Nacional de transformación digital e inteligencia artificial ●

Directrices éticas para una inteligencia artificial de confianza ●

Orientación para la regulación de la aplicación de la Inteligencia Artificial y los estándares de IA ●

Hoja de ruta de la IA ●

2010   **2012**   **2014**   **2016**   **2018**   **2020**

2011   **2013**   **2015**   **2017**   **2019**

EE.UU. - Plan de Acción Nacional de Gobierno Abierto de EE.UU. ●

Licencia de Gobierno Abierto del Reino Unido y data.gov.uk ●

Estrategia del Centro de Datos del Gobierno Australiano 2010-2025 ●

Plan Nacional Trienal de IA ●

GDPR ●

Estrategia sobre la IA ●

Plan estratégico nacional de investigación y desarrollo para la IA ●

Acuerdo sobre el sector de la IA ●

Estrategia de I+D en IA ●

Libro blanco sobre la estandarización de la inteligencia artificial ●

Agenda Nacional Mexicana de Inteligencia Artificial ●

Libro blanco sobre IA - Un enfoque europeo de la excelencia y la confianza ●

Ley de aplicación de la Carta Digital ●

Nuevo GDPR (Post Brexit) ●

In 2017-19, DPA worked with DNP and iNNPULSA on Colombia's 1ˢᵗ Big Data Strategy

Estrategia Nacional de Big Data para Colombia

National Big Data Strategy for Colombia

Documento 1: Estado del arte y análisis comparativo de estrategias nacionales de Big Data

Document 1: State of affairs and comparative analysis of national Big Data Strategies

This study provides an analysis of the main characteristics of the existing government strategies related to Big Data, while clarifying the operational concept of Big Data used throughout this project.

Descargar

Documento 2: Buenas prácticas sobre cómo los gobiernos pueden implementar una estrategia para estimular el ecosistema de negocios Big Data

Document 2: Best practices on how governments can implement a strategy to stimulate the Big Data business ecosystem

This study provides an analysis of levers and channels - from tax incentives to regulatory guidelines, legal frameworks to fiscal interventions - through which a modern government such as Colombia can stimulate this sector and its contribution to the country's economy.

Descargar

Documento 5: Propuesta de plan de implementación de la estrategia de Big Data para el Estado

Document 5: Implementation plan proposal for the Big Data Strategy of the State

This paper has four main objectives: 1) to strengthen the Big Data ecosystem in the country; 2) to develop new analytical capacities for the greater use and production of better quality information; 3) to foment the digital economy, and 4) to promote the adaptation of Artificial Intelligence solutions in public administration.

Descargar

Documento 6: Impacto económico y social de la implementación de la estrategia

Document 6: Economic and social impact of implementing the strategy

This study focuses on the potential economic and social benefits in the different areas of the Big Data ecosystem. The first part of the document discusses the limits and benefits of the Big Data ecosystem, as well as how to measure the value it generates. The second part has the objective of relating the potential economic and social benefits of the Big Data ecosystem with the proposed strategy.

Descargar

Documento 3: Diagnóstico de la situación actual de Colombia estructurado alrededor de las seis dimensiones

Document 3: Analysis and diagnostic of the current situation in Colombia structured around six dimensions

This study is framed around six specific dimensions, including: (1) current state of the normative and ethical framework, (2) organizational and institutional dimensions, (3) physical infrastructure (technology), (4) data architecture, (5) human capital and data culture, and (6) investments related to the use of data.

Descargar

Documento 4: Recomendaciones e insumos para la definición de una estrategia nacional de Big Data

Document 4: Recommendations and inputs for the definition of a national Big Data strategy

This paper provides a series of recommendations and inputs for the definition of the Big Data strategy for the Colombian Government. It aims to provide short, medium and long term actions that will enable the government to overcome the technical, social and human capital challenges that currently exist with regards to data and Big Data.

Descargar

Documento 7: Proyecto piloto de analítica de datos - Supervivencia empresarial

Document 7: Pilot project of data analysis - Business survival

This document describes the data analysis process of the pilot developed by EyS Soluciones Empresariales, who mainly used Survival Analytics to create a predictive and diagnostic data analysis model of the determinant factors for the survival and exportation of Colombian businesses.

Descargar

Documento 8: Proyecto piloto de analítica de datos - Detección de anomalías en el SISBEN y movilidad

Document 8: Pilot Project of data analysis - Detection of anomalies in the SISBEN and mobility

This paper describes the results of the pilot project, the data and methodologies used as well as the challenges and obstacles found during the process. It is divided in the four stages of the project: problem analysis, data exploration, machine learning techniques, and model evaluation.

Descargar

Documento 9: Definición de la estrategia de Big Data para el estado colombiano y para el desarrollo de la industria de Big Data en Colombia

Document 9: Definition of the Big Data strategy for the Colombian state and for the development of the Big Data industry in Colombia

This document includes the recommendations and guidelines for how the Colombian government can advance data analysis projects, providing a Big Data model for those who seek to resolve, analyze, and understand at a greater granularity the problems of public management through data. The objective is to guide the reader through the different sections and variables that have to be taken into account to formulate a public management project that positively impacts the public policies with data.

Descargar

Now, DPA and the Spanish cooperation agency AECID are collaborating on a study on the topic in LAC, including AI and focusing on 5 countries, but covering the region

Big Data para el desarrollo sostenible en Latinoamérica y el Caribe

BOLIVIA   EL SALVADOR   PERU
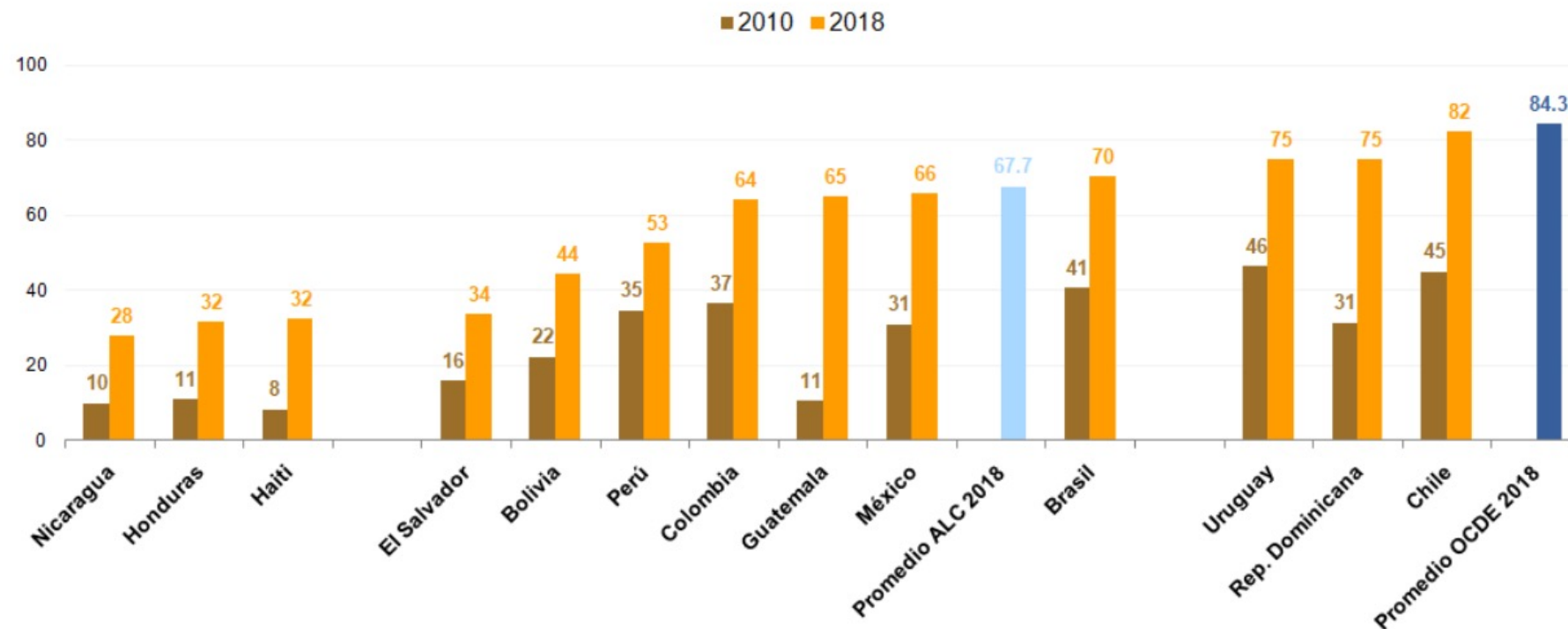
REPÚBLICA DOMINICANA   GUATEMALA

21 Julio 2021

DATA-POP ALLIANCE

aecid
Agencia Española de Cooperación Internacional para el Desarrollo

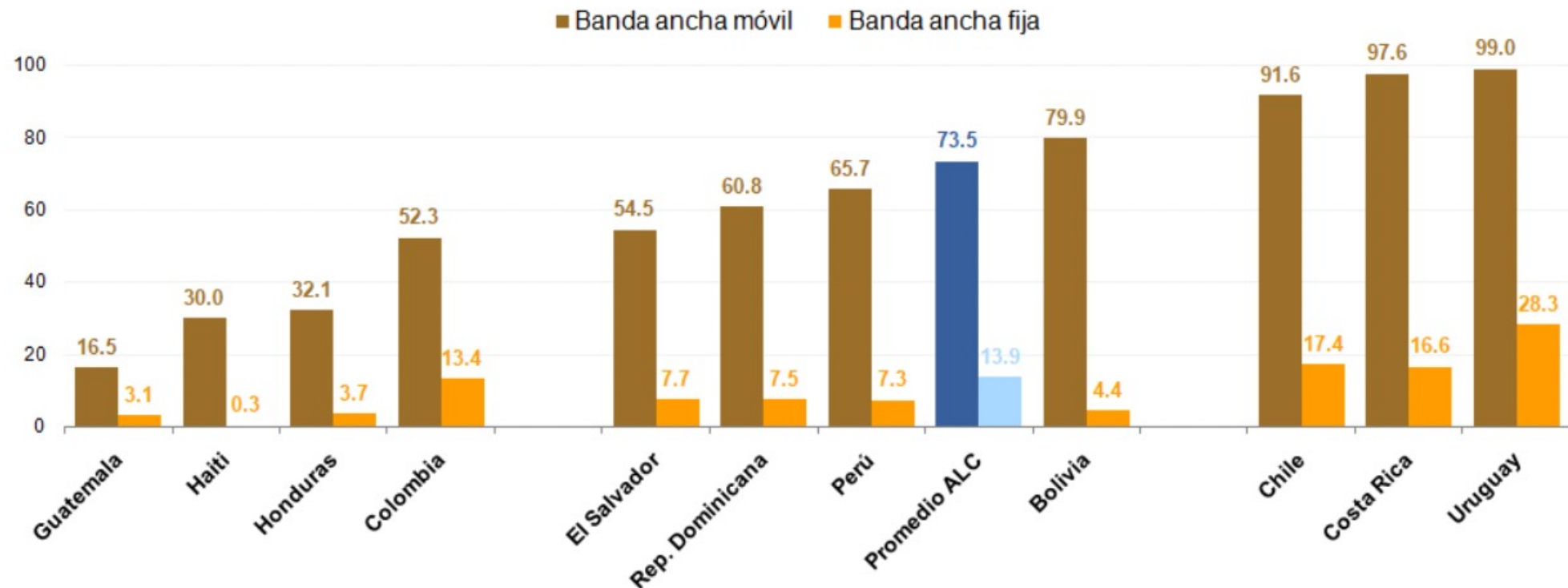# LAC countries vary quite a bit in their level of digitalization and readiness…

Gráfico 2.1. Usuarios de Internet en países seleccionados de América Latina y el Caribe, 2010 y 2018 o más reciente

■ 2010 ■ 2018

País	2010	2018
Nicaragua	10	28
Honduras	11	32
Haití	8	32
El Salvador	16	34
Bolivia	22	44
Perú	35	53
Colombia	37	64
Guatemala	11	65
México	31	66
Promedio ALC 2018		67.7
Brasil	41	70
Uruguay	46	75
Rep. Dominicana	31	75
Chile	45	82
Promedio OCDE 2018		84.3

Fuente: adaptado de OCDE et al. (2021), "Usuarios de Internet en países seleccionados de América Latina y el Caribe, 2010 y 2018 (o más reciente)", in Perspectivas económicas de América Latina 2020: Transformación digital para una mejor reconstrucción, OECD Publishing, Paris, https://doi.org/10.1787/51e1bdd0-es (consultado el 21 de agosto de 2020).
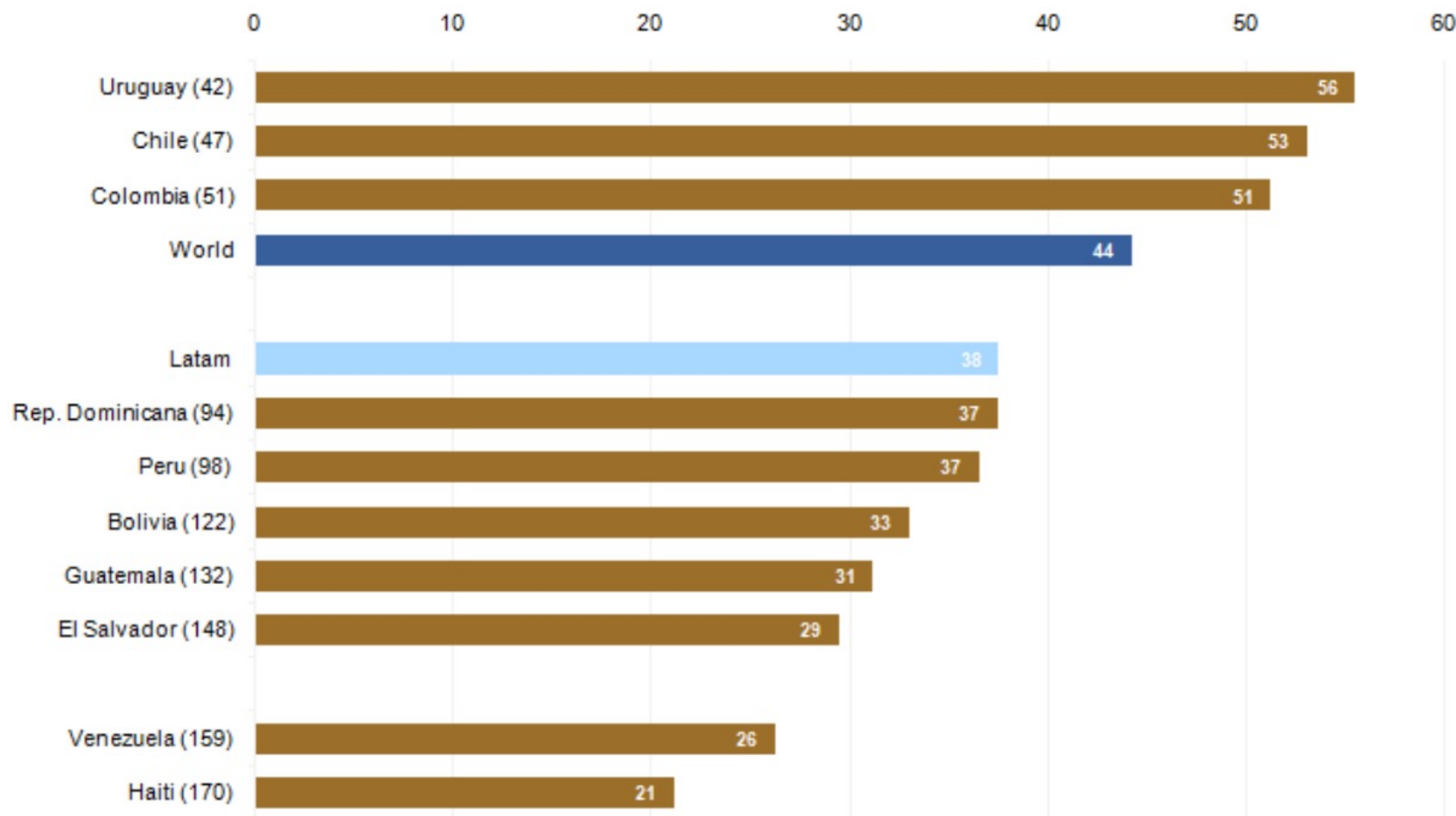
LAC countries vary quite a bit in their level of digitalization and readiness…

Gráfico 2.2. Penetración de banda ancha móvil y fija en países seleccionados de América Latina y el Caribe, 2018 o más reciente, abonos activos por cada 100 habitantes

■ Banda ancha móvil ■ Banda ancha fija

Guatemala: 16.5 / 3.1
Haití: 30.0 / 0.3
Honduras: 32.1 / 3.7
Colombia: 52.3 / 13.4
El Salvador: 54.5 / 7.7
Rep. Dominicana: 60.8 / 7.5
Perú: 65.7 / 7.3
Promedio ALC: 73.5 / 13.9
Bolivia: 79.9 / 4.4
Chile: 91.6 / 17.4
Costa Rica: 97.6 / 16.6
Uruguay: 99.0 / 28.3

Some LAC countries are more advanced

Gráfico 2.7. Government AI Readiness Index (2020) (países seleccionados)

País	Valor
Uruguay (42)	56
Chile (47)	53
Colombia (51)	51
World	44
Latam	38
Rep. Dominicana (94)	37
Peru (98)	37
Bolivia (122)	33
Guatemala (132)	31
El Salvador (148)	29
Venezuela (159)	26
Haiti (170)	21

Fuente: elaboración propia en base a datos del Índice de preparación del gobierno para la inteligencia artificial, 2020
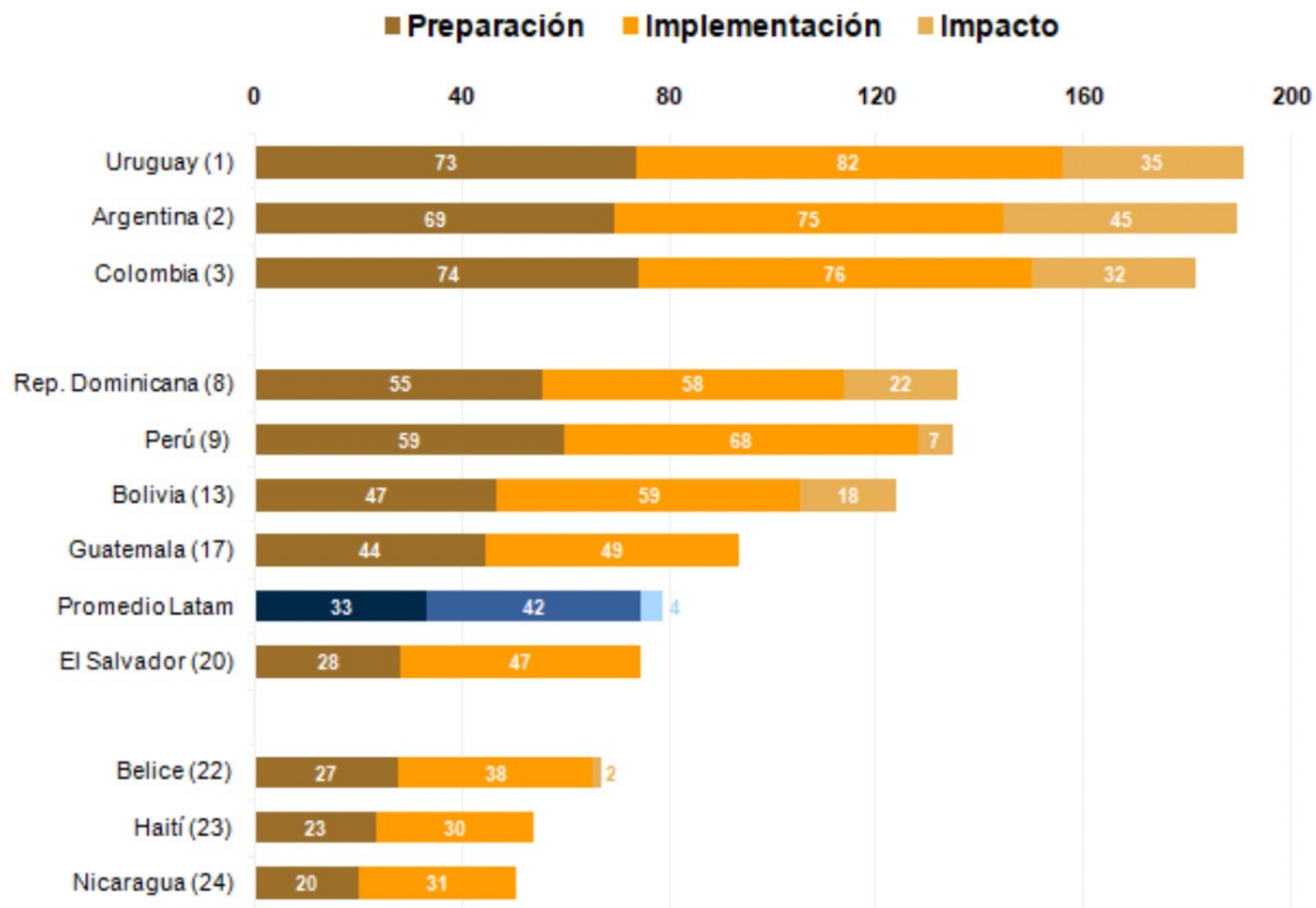
Nota: País (Ranking AI readiness)

Gráfico 2.4. Índice de Desarrollo del E-Gobierno (2020) (países seleccionados)
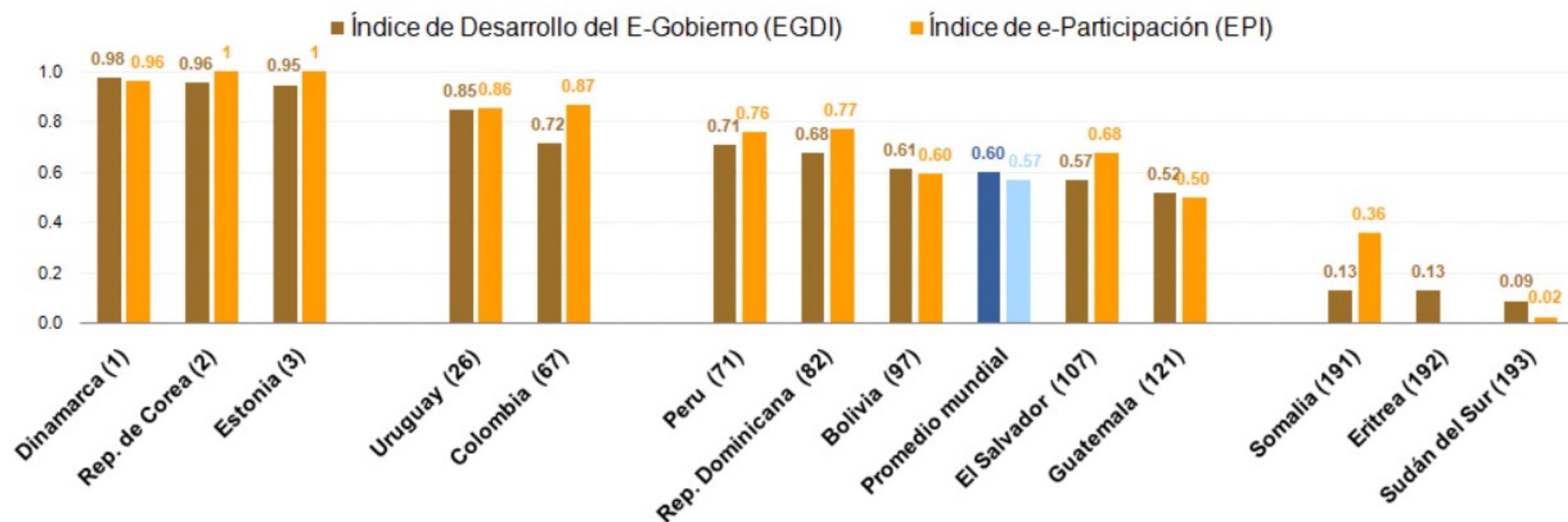
Fuente: elaboración propia en base a datos La Encuesta de Gobierno Electrónico 2020

Nota: País (Ranking EGDI)

Gráfico 2.6. Barómetro de Datos Abiertos América Latina y el Caribe (2020) (países seleccionados)
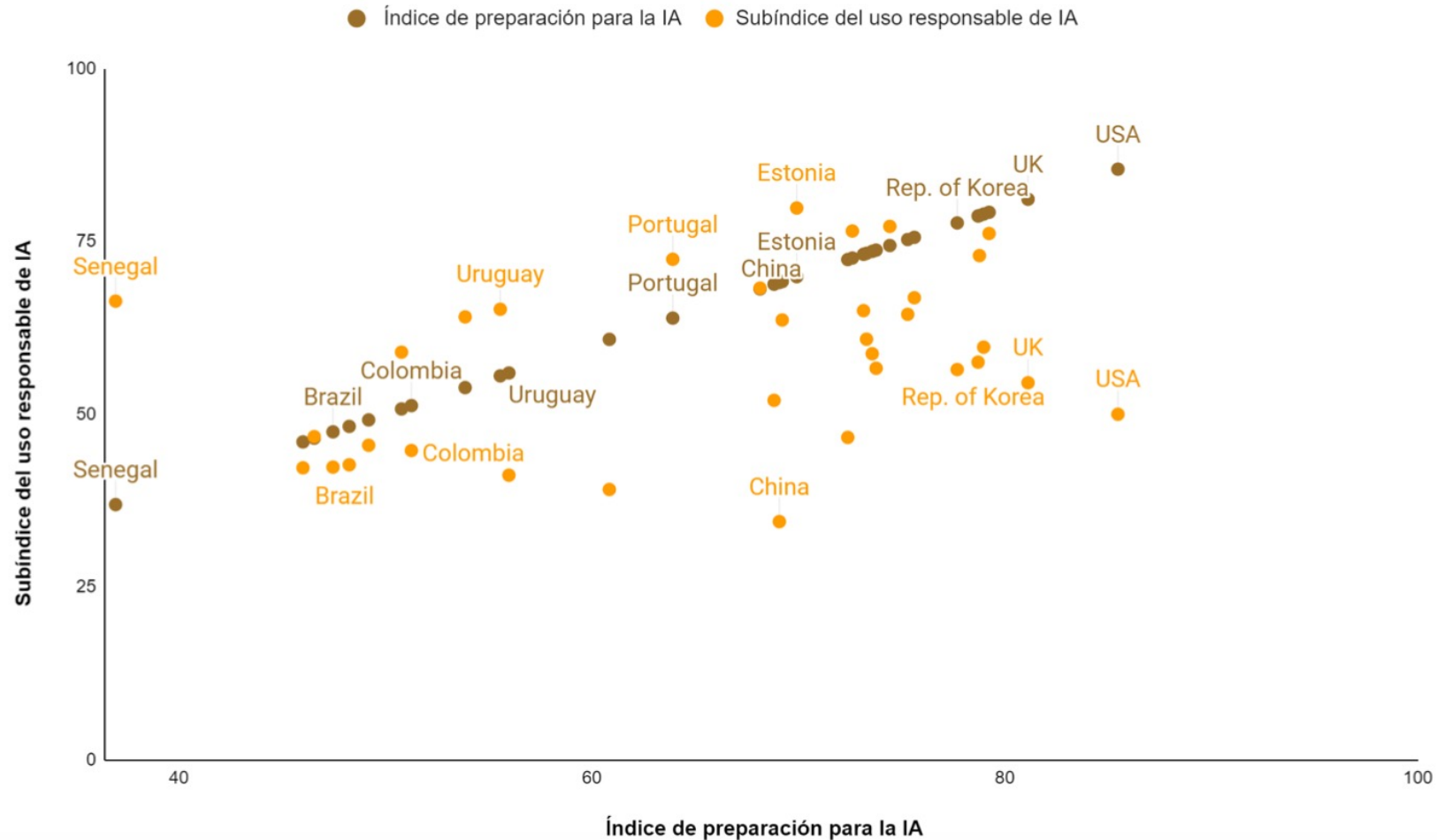

Legend: ■ Preparación ■ Implementación ■ Impacto

País	Preparación	Implementación	Impacto
Uruguay (1)	73	82	35
Argentina (2)	69	75	45
Colombia (3)	74	76	32
Rep. Dominicana (8)	55	58	22
Perú (9)	59	68	7
Bolivia (13)	47	59	18
Guatemala (17)	44	49	
Promedio Latam	33	42	4
El Salvador (20)	28	47	
Belice (22)	27	38	2
Haití (23)	23	30	
Nicaragua (24)	20	31	

Fuente: elaboración propia en base a datos del Barómetro de Datos Abiertos América Latina y el Caribe, 2020
Nota: País (Ranking ODB).

But even they are somewhat lagging behind

Gráfico 1.1. Comparación del Índice de preparación del gobierno para la inteligencia artificial y el subíndice del uso responsable de la inteligencia artificial (2020)

Different countries have (or don't have) different policies

Figura 2.4. Principales instituciones en las estrategias de big data y digitalización - países foco

Bolivia

Agencia de Gobierno Electrónico y Tecnologías de Información y Comunicación (AGETIC)

Consejo para las Tecnologías de Información y Comunicación del Estado Plurinacional de Bolivia (CTIC-EPB)

El Salvador

Secretaría de Innovación

Guatemala

Ministerio de Gobernación

La Secretaría Nacional de Ciencia y Tecnología (SENACYT)

Comisión Presidencial de Gobierno Abierto y Electrónico (GAE).

Secretaría de Planificación y Programación de la Presidencia

(SEGEPLAN)

Perú

Secretaría de Gobierno Digital

CODESI

Secretaría de Gestión Pública

República Dominicana

Oficina Presidencial de Tecnologías de la Información y Comunicación (OPTIC)

Dirección General de Ética e Integridad Gubernamental (DIGEIG)

Instituto Dominicano de Telecomunicaciones (INDOTEL)

Nota: Selección de instituciones que juegan un papel protagónico en el desarrollo de la transformación digital y políticas de datos a nivel nacional.
Fuente: Elaboración propia.

Figura 2.1. Intensidad de las dimensiones digitales en los planes nacionales de desarrollo, por trampa del desarrollo, países seleccionados de América Latina y el Caribe (2019)

País	Total	Productividad	Institucional	Vulnerabilidad social	Ambiental
Bolivia					
Brasil					
Colombia					
República Dominicana					
El Salvador					
Guatemala					
Perú					
Uruguay					

Fuente: Adaptado de Perspectivas económicas de América Latina 2020, OECD 2019 **Nota:** Un color más oscuro indica una mayor intensidad de la transformación digital. Para configurar la tabla, se vinculó cada política de los planes nacionales de desarrollo con una 'trampa del desarrollo' y, posteriormente, se calculó la frecuencia relativa de las palabras "computacional", "digital", "digitalización", "electrónico", "informática", "inteligente", "Internet", "online", "tecnología", "tecnológico", "virtual" y todos sus derivados con respecto a cada trampa.

Human capacities are critical, and overall missing…..

Gráfico 3.1. Índice de Competitividad del Talento Global (Informe 2020) - Puntuación GTCI

País	Puntuación
Switzerland (1)	81
USA (2)	79
Singapore (3)	78
Promedio Mundial	50
Uruguay (51)	47
Colombia (74)	40
Perú (77)	39
Rep. Dominicana (81)	37
Guatemala (95)	35
Bolivia (107)	30
El Salvador (110)	30
Congo (130)	18
Angola (131)	16
Yemen (132)	12
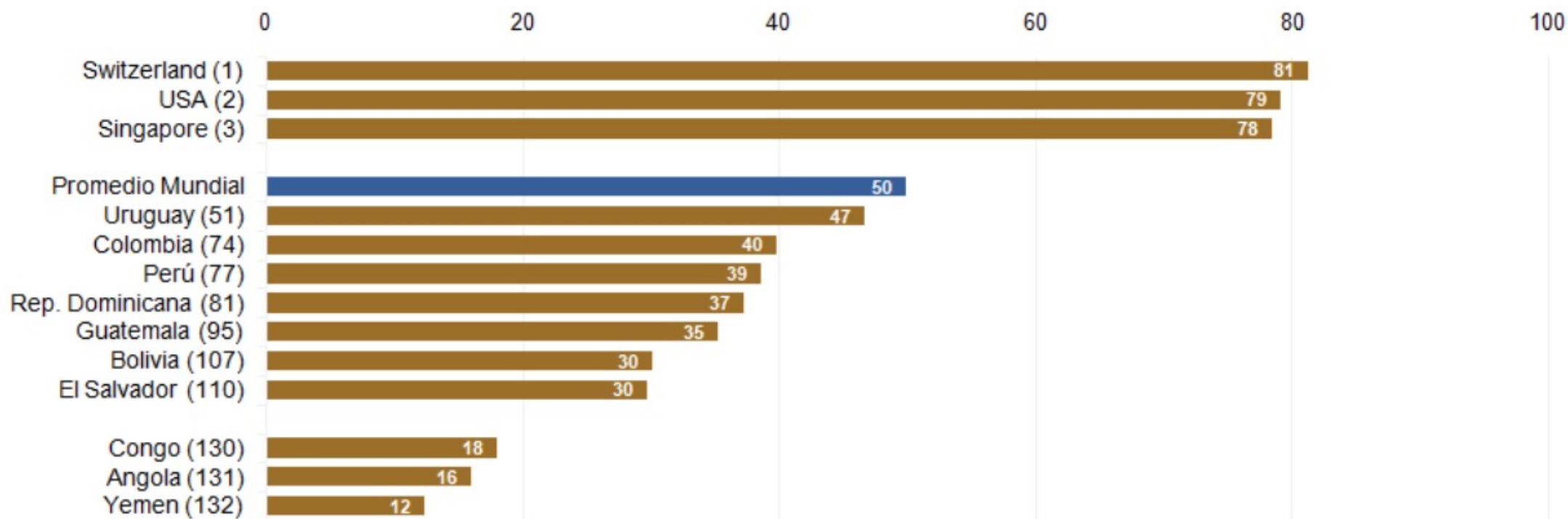
Fuente: Elaboración propia en base a INSEAD (2020): The Global Talent Competitiveness Index 2020: Global Talent in the Age of Artificial Intelligence, Fontainebleau, France. **Nota**: valor entre paréntesis denota ranking mundial.

Keys: "data literacy", data strategies, and "rational compassion". *Data as 21st century global language*

C. Lévi-Strauss
de l'Académie française
Tristes tropiques

MOBILISER LE
BIG DATA
POUR LE
DÉVELOPPEMENT
DURABLE
Atelier de formation professionnelle

GUIDE DES PARTICIPANT-ES

Nations Unies DAKAR
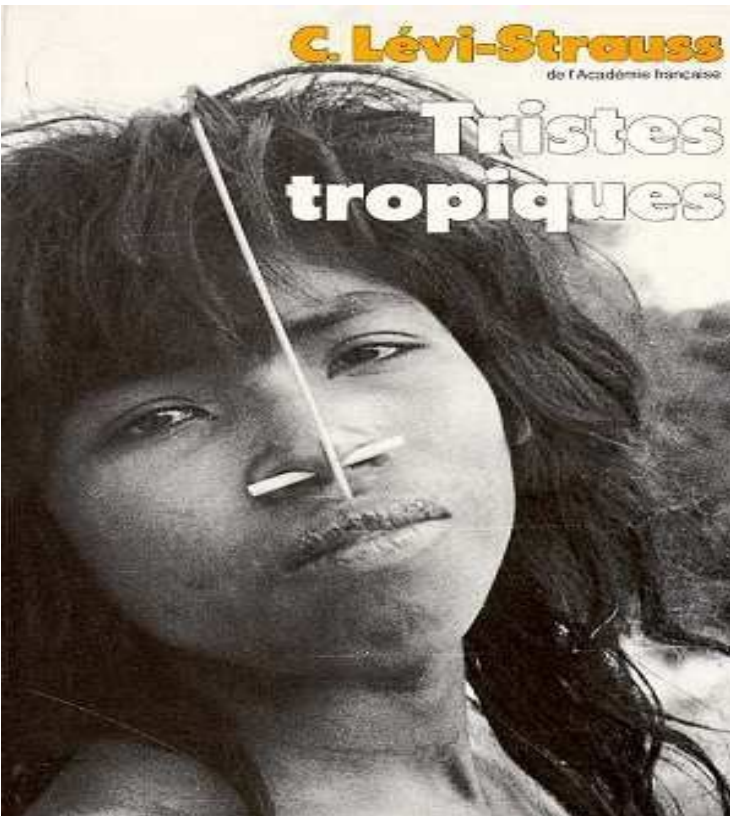
Building Literacy for the Data Generation

December 18, 2015

A unique opportunity exists to develop data literacy education for children born into a world shaped by big data.

Beyond Data Literacy:
DATA-POP ALLIANCE Reinventing Community
WHITE PAPER SERIES Engagement and Empowerment

The real digital divide is between families that limit screen time and those that don't

Naomi Schaefer Riley, Contributor

EMPODERADATA:
DATA LITERACY
ASSESSMENT
AND
SUSTAINABLE
DEVELOPMENT GOALS
DATA GAPS

Brazil, Colombia and Mexico

Vanessa Higgins
Valentina Casasbuenas
Julie Ricard
Jackie Carter

October 2019

MANCHESTER 1824
The University of Manchester

DATA-POP ALLIANCE

Blockchain and its application to collective challenges

Although it has most commonly been associated with cryptocurrency, since its release in 2008 with the white paper "Bitcoin: A Peer-to-Peer Electronic Cash System", the blockchain system has been always praised for its potential to transform various sectors. Namely, blockchain can be used to address environmental challenges, enable e-voting, expand digital mobile ticketing, facilitate health care service and even contribute to achieving the Sustainable Development Goals (SDGs). One of the biggest appeals of this system is its intrinsic characteristics of transparency. According to Kazuhiro Gomi (CEO of NTT Research and leading researcher in physics, informatics, cryptography and information security) the fact that blockchain technology automatically includes a public record of any data transactions that have been conducted, eliminates the possibility of manipulating the system without other users noticing and being alerted. This is precisely what makes blockchain so appealing for other sectors to adapt into their needs of privacy and security.

INNOVATION

How the Blockchain Brings Social Benefits to Emerging Economies

Nov 28, 2018

📍 Africa, Global Focus, India, North America

Developing countries such as India, Kenya and others in East Africa are discovering an increasing array of applications for blockchain, the decentralized ledger technology that promises a secure, peer-to-peer mechanism for verifying information. Blockchain is finding innovative uses in banking and financial services, supply chains, agriculture and in managing land ownership records (land titling) in those countries, according to panelists who spoke at the Wharton India Economic Forum held recently in Philadelphia.

However, many laws in both developing and developed countries have not kept pace with digital advancements, and they continue to require paper-based documentation, preventing participants from taking full advantage of the technology, they said. Although a decade has passed since the blockchain's invention, its technology is still evolving and being tested.

TSG CONSUMER PARTNERS

Supports K@W's Innovation Content

https://knowledge.wharton.upenn.edu/article/blockchain-brings-social-benefits-emerging-economies/

Key Recommendations

1. Think and act with a system's / ecosystem's approach; it is not just about data and tech
2. In the region, some countries can serve as examples of what to do (and what not to do) and partners – collaborate, exchange!
3. Using the 3Cs or 4Cs framework can help identify and fix strengths and gaps
   1. Crumbs: facilitate the use and sharing of relevant data
   2. Capacities: invest in infrastructure and human capacities
   3. Communities: develop partnerships and enabling regulations
   4. Culture! address cultural requirements and resistance

We need **new technological, governance, and cultural standards for future human generations to be "Human AI" learning societies**, in LatAm and globally

NEWS FEATURE · 29 MAY 2019

Can tracking people through phone-call data improve lives?

Researchers have analysed anonymized phone records of tens of millions of people in low-income countries. Critics question whether the benefits outweigh the risks.

Amy Maxmen

...k off us," Rivers explains.

Letouzé, de Montjoye and their colleagues [ar]e piloting a system called Open Algorithms [(O]PAL) in Senegal and Colombia. As well as [ru]nning analyses on phone-company servers, [th]eir model includes a committee that vets and [sh]apes researchers' questions so that the data [an]alysed are less specific. For instance, if aid [w]orkers want to know how many people leave [Se]negal's capital city Dakar each week, the com[m]ittee can decide that records should be aggre[ga]ted by day, rather than by hour. This reduces [th]e number of extra, unapproved questions that the results can answer. "It's not a perfect system," de Montjoye says, "but we are trying to find a way to mitigate risks, while making sure data can be used for good."

...is becoming harder to sustain in villages." Global funding for malaria has plateaued in the past few years, she points out — and with it, progress.

The same practical argument could be made against research on parasite genetics. But Nyunt says that call-record analyses trouble her more, because people haven't consented to take part.

"NOW IS THE TIME TO PUT IN PLACE STANDARDS TO DO THIS SAFELY, AT SCALE AND ETHICALLY."

DATA FOR DEVELOPMENT
In 2012, the mobile-phone company Orange, together with data scientists at the UN and several universities, held a 'Data for Development' challenge to encourage researchers to explore positive uses for call-detail records. Phone companies mostly analyse the records to boost their businesses, says Robert Kirkpatrick, director of UN Global Pulse, an initiative to harness big data. "We wanted to show how it could be used for the public good," he says.

Orange let scientists analyse anonymized call records from customers in Côte d'Ivoire. In one project, researchers found that brief calls surged before small violent events in Côte d'Ivoire, and suggested that future analyses could help officials to predict danger and thus intervene — but that idea hasn't been taken up.

...erlands, defends the project's worth. Anyone who might want to harm any of the 3.6 million Syrian refugees in Turkey already knows their neighbourhoods, he argues. But call-record intelligence might help policymakers by giving them quantitative information about refugee movements. And an ethics committee vetted the results: when research indicated refugees were working at a location illegally, for example, the committee told them not to publish the finding.

Responding to the charge that such data challenges have not helped people, Kirkpatrick says exploration was a necessary first step. The next phase in call-records research, he says, should be cost–benefit analyses that look at the investment needed to conduct a study, roll out an intervention and appraise the advantages for communities.

SECURITY AND CONSENT
In the meantime, exploratory studies continue.

Since last year, groups including Flowminder and phone companies that are headquartered in Europe must comply with the European Union's general data-protection regulation. Although anonymized and aggregated data seem to be exempt, Letouzé thinks that the law signals a trend towards privacy, and suggests that data scientists should consider how they might incorporate consent into their studies. OPAL is planning to send subscribers a text message asking if they want to opt out, which causes Letouzé some concern. "There are studies showing that when you give people an option, you lose about half," he says. He'd like to change that by convincing people of the worth of their studies, and by giving them assurances about data security.

UNINTENDED CONSEQUENCES
Advocates for data security and human rights say that, although technical changes are welcome, more careful risk assessments are required, because records don't need to be...

Gracias!

eletouze@datapopalliance.org
@mit.edu