How Big Data, Open Algorithms and Artificial Intelligence Can Drive Smart Cities and Societies: Towards Human AI Ecologies

Emmanuel Letouzé, PhD
Director, Data-Pop Alliance | Program Director, OPAL Project
Visiting Scholar, MIT Media Lab | Connection Science Fellow, MIT

SmartStatistics4SmartCities Seminar
Kalamata, Greece, Oct 5-6 2018
Trust me

I'm a data scientist
• Fake news. Biases. Automation. Echo chambers. Information overload. CO₂ emissions...

• (Big) Data is getting a bad name. Are data, algorithms and AI threats to sustainable development and democracy?

• Can we instead envision and build a world where Big Data, Open Algorithms and AI drive better, fairer, more sustainable and more resilient cities and societies?

• Let’s call these ”Human Artificial Intelligence ” ecologies. What would it look like, and take?
How do(es) AI(s) Work?

Is this a city or a beach?

1. Try to guess / recognize. Right or Wrong?
2. Correct: +1. Reward!
3. Incorrect: -1. Penalty!
4. Repeat and learn through many feedback loops.

⇒ (The) machine (is) learning
Artificial intelligence is the simulation of human intelligence processes by computer systems, especially artificial neural networks (ANNs) inspired by the biological neural networks that constitute animal brains, which can "learn" (i.e. progressively improve performance on) through iterations and feedback. AIs are powered by algorithms that learn to automate parts or all of tasks, and the machines they power. (It’s also what has not been invented yet).
Is AI some new (black) magic? *No...but...*

*No...*
1. It is **at least 60+ years old**.
2. It still generalizes poorly. It has no sense of context. It is still pretty stupid.
3. We are **far from general AI**.
4. Humans are still in control (for better or worse)

*...but...*
1. The **(good) magic / core** of the current AI is the credit assignment function to encourage and reinforce neurons / functions that help the most achieve the goal (and reverse if not)
2. The key difference and is data. **Big Data.**
Some Early Applications

**Scientific Prize and Ethics Mention:** Construction of socio-demographic indicators with digital breadcrumbs

F. Bruckschen (1), T. Schmid (2), T. Zbiranski (1)

We show that socio-demographic indicators such as population, age, literacy, poverty, religion, ethnicity, electricity supply and others can be estimated in unprecedented detail and virtually ad-hoc using antenna to antenna traffic data only. We offer a uniform approach that can be easily extended to other variables. Results are tested for spatio-temporal robustness and visualized as heat maps.

(1) Humboldt Universität Berlin, Germany - (2) Freie Universität Berlin, Germany

---

Home to Clinics

Commute in México

Source: Noriega, Pentland

Source: Letouzé, 2013

Source: Letouzé, 2014
MIT Prof Alex ‘Sandy’ Pentland:

“The big question that I'm asking myself these days is how can we make a human artificial intelligence? (...) I don't want to think small—people talk about robots and stuff—I want this to be global. (...) What would happen if you had a network of people where you could reinforce the ones that were helping and maybe discourage the ones that weren't? That begins to sound like a society or a company”.

The Human Strategy. www.thehumanstrategy.mit.edu
## Vision of a “Human AI”

<table>
<thead>
<tr>
<th>1. Key principle</th>
<th>2. Key features</th>
<th>3. Key requirements</th>
</tr>
</thead>
</table>
| Taking the **key insights of AI especially**  
  • role of data  
  • credit assignment function reinforcing “neurons” that work (teams, groups, policies) through learning  
  + applying this general framework to entire societies | Leveraging **human-machine complementarities**:  
  • humans do the strategy and oversight and machines do the tactics and bookkeeping  
  • Humans + Machines >> Humans or Machines  
  • New jobs will be created (e.g. machine prison guards but more social workers)  
  • Resulting ecologies are more agile and resilient | **Good data** on the system’s functioning and performance  
  • Good **feedback and response systems** (i.e. “human or society in the loop”)  
  • Some general **agreement on inputs (facts) and outputs (goals)**  
  • Sufficient **human skills and trust** to oversee, implement, learn, adapt, and again |
Main Challenges to a Human AI

1. Some powerful agents have strong incentives for this not to work (e.g. economic and political monopolies benefit from status quo)
2. Most societies / countries currently lack the appropriate data connections, capacities, and culture for this
3. There is widespread (and growing?) digital and analog segregation with distrust, disdain, echo chambers, alternative facts narratives, hampering cooperation and consensus building
4. We know AI can and has been used to nurture 3. (cf Facebook newsfeed; Amazon Prime..)
“Open Algorithms”: A Bold New Vision and Project

Open algorithms: A new paradigm for using private data for social good

By Thomas Roca, Emmanuel Letouzé | 18 July 2016

The Open Algorithm project: Developing Indicators, capacity and trust

To address the complex challenge of data access, Orange, MIT Media Lab, Data-Pop Alliance, Imperial College London and the World Economic Forum — supported by Agence Française de Développement and the World Bank — are developing a platform to unleash the power of “big data” held by private companies for public good in a privacy-preserving, commercially sensible, stable, scalable and sustainable manner.

Mette le Big Data privé au service du bien public

Le projet Open Algorithm vise à utiliser les données d'entreprises privées pour des actions de développement.

UN Big Data for Official Statistics Conference
Bogotá, Nov. 8 2017
OPAL: 1st Generation Data Systems and Standards

1. Partner private companies (here a telecom operator) allow OPAL to access its servers through a secured platform. The data never leave the servers.

2. Certified open algorithms developed by developers are sent and run on the servers of partner private companies, behind their firewalls.

3. A governance system including a Council for the Orientations of Development and Ethics (CODE) ensures that the algorithms and use cases are ethically sound, context relevant, etc.; users benefit from capacity building activities.

4. Key indicators derived from private sector data such as population density, poverty levels, or mobility patterns, feed into use cases in various public policy and economic domains. Data are safe, minimized, used (more) ethically.
OPAL Started with 2 pilots in Colombia and Senegal with 2 Major Telcos and their NSOs
Key to all this: Building Capacities and Connections

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.

The question of how growing up with digital technology shapes a generation’s outlook has fueled discussion since the description “Digital Natives” was coined in 2001. As commentators begin to weigh in on the experience of those born in the decade-and-a-half since then, Emmanuel Letouzé, director of Data-Pop Alliance, believes one milestone merits special consideration: The advent of Big Data.

"Writing is a strange thing. If my hypothesis is correct, the primary function of writing, as a means of communication, is to facilitate the enslavement of other human beings”.

The fight against illiteracy goes on par with an increase in the control of the Power over citizens.”

Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data

October 2015

“We define data literacy as “literacy in the age of data”, i.e. “the desire and ability to constructively engage in society through or about data”.

DATA-POP ALLIANCE
WHITE PAPER SERIES

LEVERAGING BIG DATA for SUSTAINABLE DEVELOPMENT
Professional Training Workshop

June 27-29, 2017
Mathematics Office of Nairobi (KNUC) Nairobi, Kenya
Thank you

eletouze@opalproject.org
@mit.edu
@datapopalliance.org