

# The epochal shift from Big Science to Data Science: A critical perspective

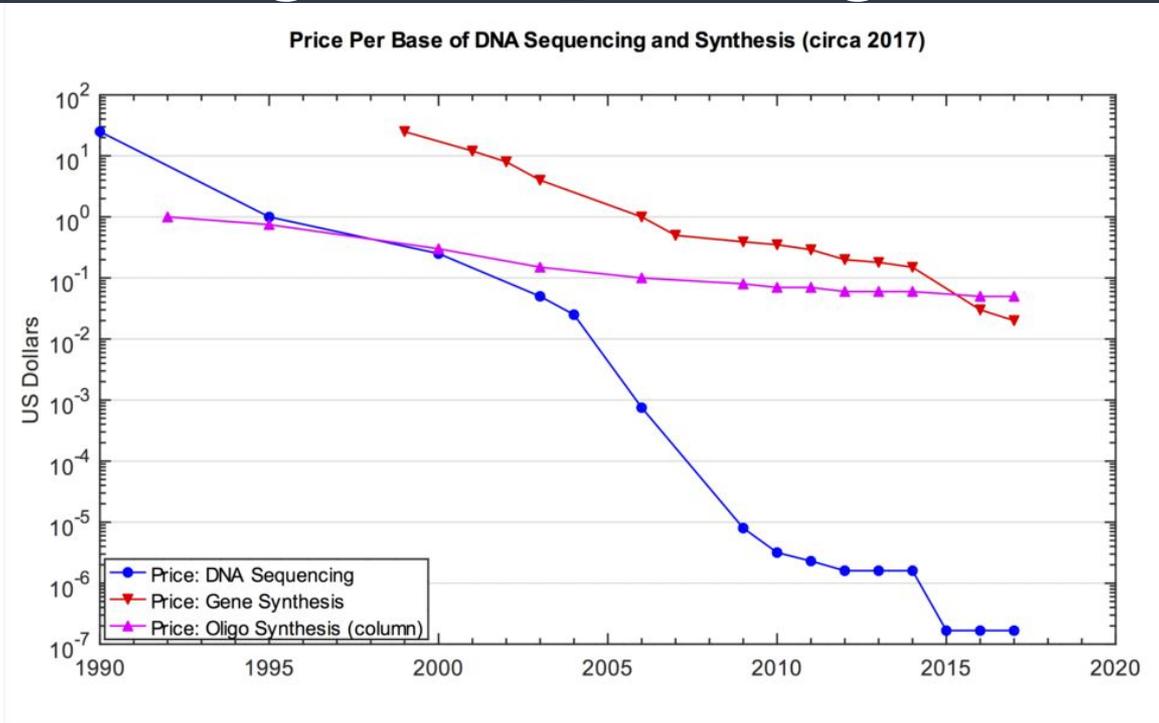
Anita Gurumurthy



Digital Sequence Information: Taking Stock and Moving Forward

African Centre for Biodiversity | Third World Network | 30 October 2020

# The digital and data revolution has brought Moore's law to gene sequencing



## A confluence of breakthroughs in biological science, with the development of computing, automation, and AI, are fueling a new wave of innovation.

### 4 arenas of biological innovation



| Definitions                    | <b>Biomolecules</b>  | <b>Biosystems</b>   | <b>Biomachine interfaces</b>  | <b>Biocomputing</b>  |
|--------------------------------|--|---|---|--|
| <b>Mapping</b>                 | Cellular processes and functions via measuring intracellular molecules (eg, DNA, RNA, proteins) in the study of "omics" <sup>2</sup> | Complex biological organizations and processes and interactions among cells     | The structure and function of nervous systems of living organisms                         | Intracellular pathways or networks of cells to return outputs based on specific conditions (for computation) |
| <b>Engineering<sup>1</sup></b> | Intracellular molecules (eg, via genome editing)   | Cells, tissues and organs, including stem-cell technologies and transplantation | Hybrid systems that connect nervous systems of living organisms to machines               | Cells and cellular components for computational processes (storing, retrieving, processing data)             |
| <b>Examples</b>                | Gene therapy for monogenic diseases  | Cultured meat grown in a lab  | Neuroprosthetics for motor control (implant or external headset) of human or robotic limb | Data storage in strands of DNA   |

<sup>1</sup>Design, de novo synthesis, or modification.

<sup>2</sup>"Omics" refers to technologies that allow the identification and quantification of molecules of a biological system.

Source: McKinsey Global Institute analysis

The digital  
bio-revolution  
could have an  
annual economic  
impact of \$ 2-4  
trillion in the next  
two decades

McKinsey (2020) has observed:

more than half of the direct impact of these applications is likely to be outside health, primarily in agriculture and consumer products

Sean Ward, Chief  
Technology Officer  
of Synthace, a  
synthetic biology  
company

As working with the physical world is becoming increasingly digital, every company that is out there is discovering that they either are a technology business or they are dead. And that is what is happening with biology: it is becoming a technology business”

# Digital genomics is now a multi-million dollar industry

The global digital genome market size crossed \$26 billion in 2018 and is projected to register over 10% CAGR up to 2025 (Global Market Insights 2019)

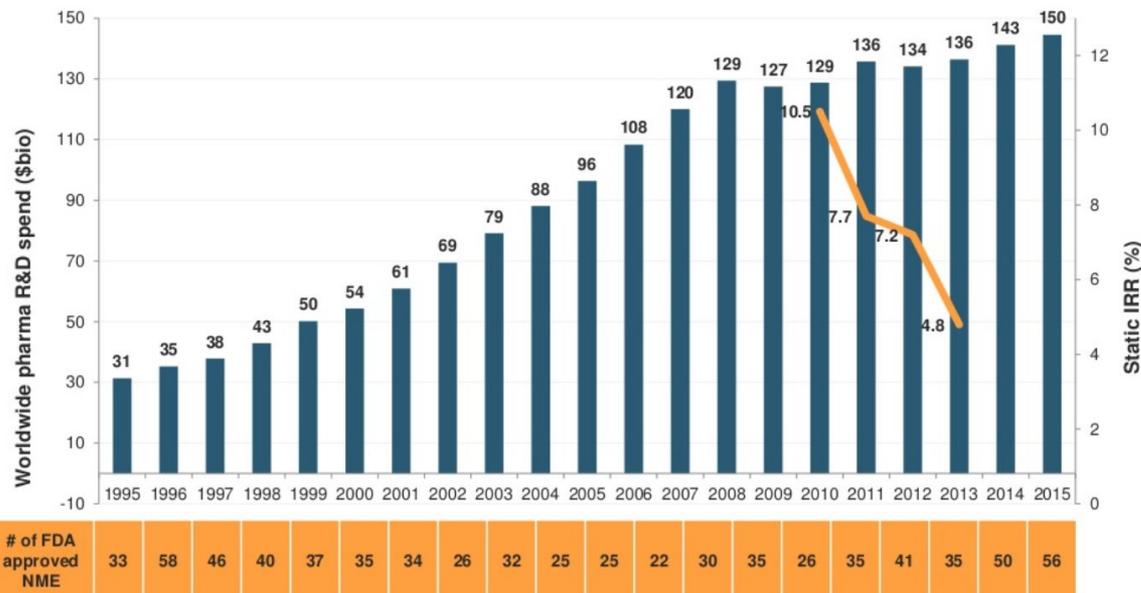
Global digital genome market and its key applications

- Diagnostics
- Agriculture & animal research
- Personalized medicine
- Drug discovery
- Treatment analysis

# Big Pharma is ramping up R&D expenditure

Pharma today is investing 5 times as much in R&D as it was two decades ago

Trends in R&D expenditure, static IRR, and NME approved by the FDA



This trend has continued.

In 2019, industry analysts found that R & D spending among the top 11 pharma companies reached a total of \$77.6 billion.

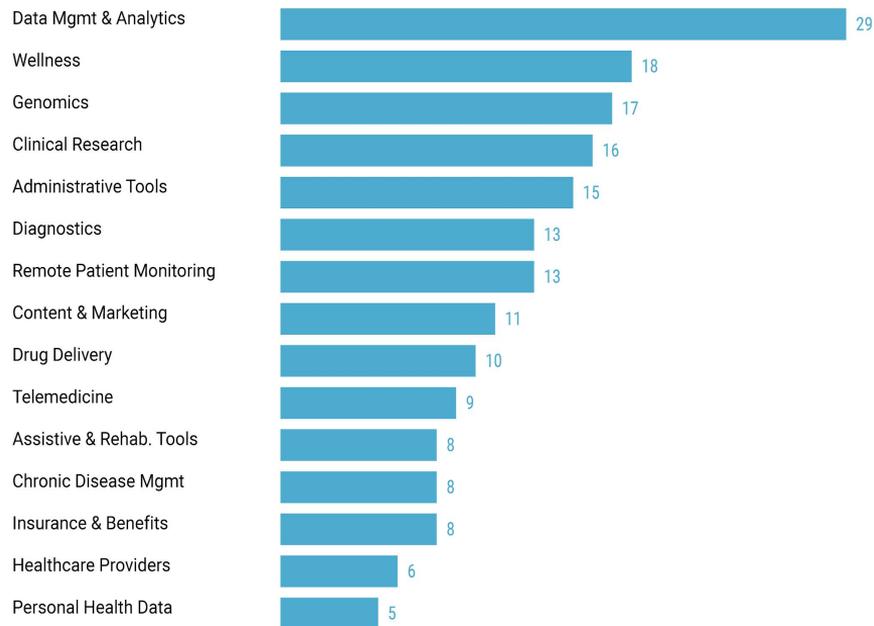
# Where is Big Pharma's health tech R & D investment going?

All of the 10 so-called Big Pharma companies (namely Novartis, Roche, Pfizer, Merck, AstraZeneca, GlaxoSmithKline, Sanofi, Abbvie, Bristol-Myers Squibb and Johnson & Johnson) have either expressly collaborated with or acquired Artificial Intelligence technologies

# Big Tech is ramping up investments in genomics

## BIG TECH FOCUSES ON DATA ANALYTICS, WELLNESS, & GENOMICS

Number of digital health startups in each category that have received funding from a big tech company since 2010.



# Big Agriculture and digital gene sequencing

The Div-Seek expose by Edward Hammond in 2016 - gene sequences and patent rights to climate change genes offered to Syngenta and DuPont Exchange

Gates Foundation now funding DivSeek!

“Today, biopiracy is carried out through the convergence of information technology and biotechnology. It is done by taking patents by ‘mapping’ genomes and genome sequences... DivSeek is a global project launched in 2015 to map the genetic data of the peasant diversity of seeds held in gene banks. It robs the peasants of their seeds and knowledge, it robs the seed of its integrity and diversity, its evolutionary history, its link to the soil and reduces it to ‘code’. It is an extractive project to ‘mine’ the data in the seed to ‘censor’ out the commons.”

**Vandana  
Shiva**

# The unmistakable materiality of digital sequence information

DSI is a neologism that originally combined the phrases “digital sequences” with “genetic sequence information”, that were used in the 2015 meeting of the CBD’s Synthetic Biology AHTEG (TWN 2019)

As we saw in COP 13, and the various contact group sessions around this, the debate was focused on whether **digital sequence: physical sequence** is a relationship like **hard copy: e-book** or **image of a corn: a corn**

But as we can learn from the data capitalism debates, this is a moot point. The question is this: even in digital digital databases, DNA sequences can be made to play material/expressive roles from which value can be generated or captured (Scott 2020)

# The pandemic moment and DSI

30th September 2020 -- announcement of a  
communique from Gates Foundation - 15  
company partnership for drug discovery.

False claim about equitable global access

Refusal to engage in open licensing and  
technology transfer

Foundation is entering into higher-profit,  
preferential advance purchase agreements with  
rich countries

# The pandemic moment and DSI (contd.)

Issues raised by 400 CSOs in April 2020, about access and benefits sharing in the COVID moment, remain unaddressed, particularly:

The demand to ensure IP rights regimes do not compromise open access sharing of technology and knowledge including technical specifications, designs, blueprints and any other know-how to scale-up local/regional manufacturing of medical products required for COVID-19 response including diagnostics, therapeutics and vaccines [including DSI of virus gene]

Are we repeating Regeneron and Gene Bank (Ebola -2014) all over again?

# Insights from the data justice debate

Fetishization of open access DSI databases akin to the free data flows dogmas in global data governance debates

DSI is not completely non-rivalrous in its material/expressive uses and there will be value capture, just like in the case of data processing for insight generation. DSI is a system resource.

Rules of access and use essential, open access will produce the Hardinian tragedy.

Benefit sharing of DSI -- community data frameworks may offer a way forward -- especially the concept of trusteeship (as evidenced by the Costa Rica model).

Big Pharma-Big Tech hybrids here to stay.

**We need radical action - NOW!**