

3rd International Systems Biomedicine Symposium
“Big Data in Health Care”
28-29 October, 2015, Luxembourg

Big Data in Healthcare: What impact over the next decade?

Dr David Brown, PhD, FRSC, FRSM

Alchemy Biomedical Consulting
Cambridge, UK
alchemybiomedicalconsulting.vpweb.co.uk
email: davidbrown1000@btinternet.com
+44 (0) 7766 686 345

The massive opportunity

- **We are in the early stages of the 3rd Industrial Revolution**

1st Industrial Revolution:
1780 -1830

2nd Industrial revolution:
1890 -1930

3rd Industrial Revolution:
mid-1990's - 2040s?



1. What drives an Industrial Revolution?

Three breakthroughs...

1. A new communication system
2. A new energy source
3. A new financial system

Breakthroughs that drove the 1st and 2nd Industrial Revolutions

1st Industrial Revolution

1770-1830, UK

1. **Energy**: coal, steam power

which drove...

2. **Communication**: canals, railways,
mass printing/education

all funded by...

3. **Financial**: London stock market

Jonathan's Coffee-House (1600s) → "Stock
Exchange" in Sweeting's Alley (1801).

2nd Industrial Revolution

1870-1914, UK/ Germany/ USA

1. **Energy**: oil, electricity

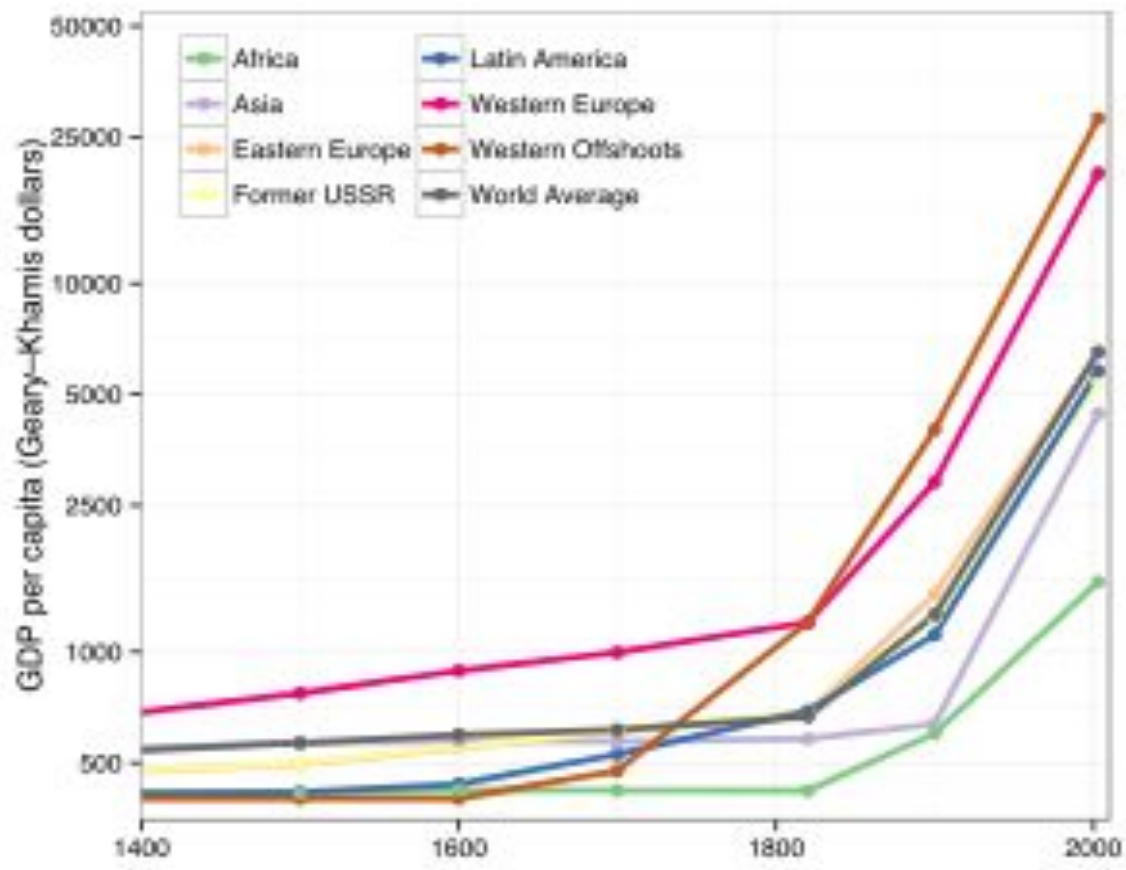
which drove...

2. **Communication**: cars, highways,
telegraph, telephone, aircraft

all funded by...

3. **Financial**: Limited liability
corporation → large scale industry

GDP per capita lift-off

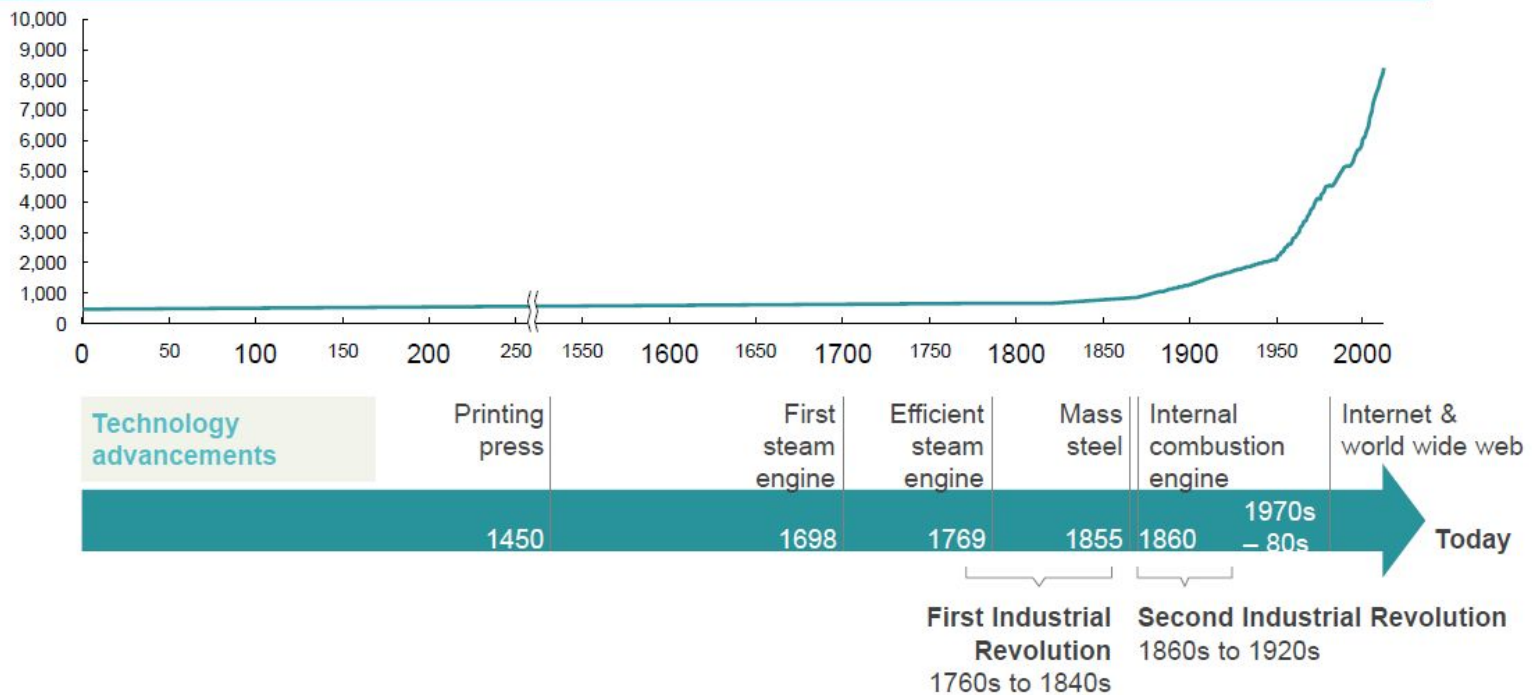


Before the 1st Industrial Revolution there was little growth in global wealth from century to century!

Growth take-off

Source: Jacques Bughins, McKinsey Global Institute, 2013

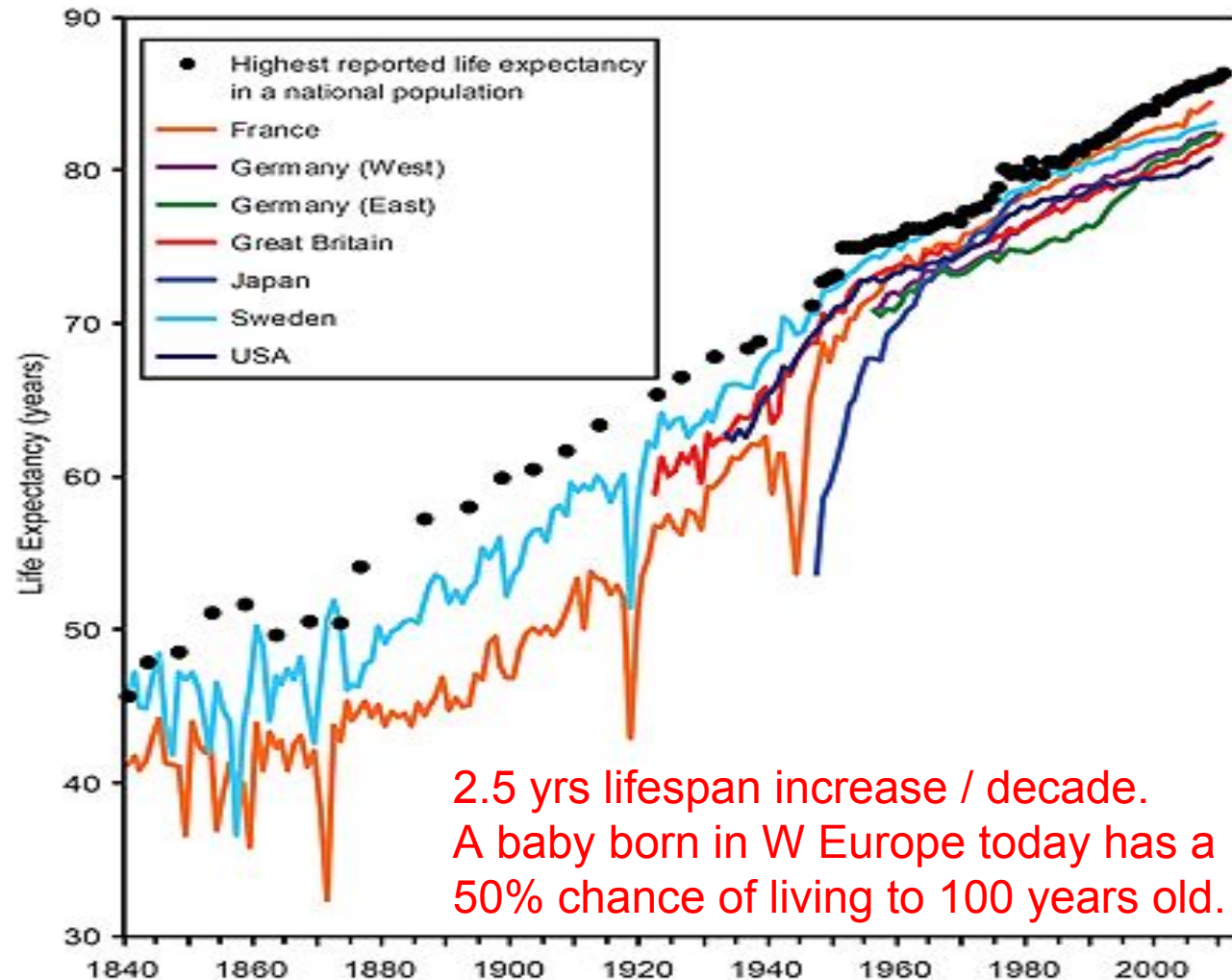
GDP per world capita
In real USD



SOURCE: Angus Maddison's "World Population, GDP and Per Capita GDP, 1-2003 AD"; Projection based on Global Insight economic data; WIPO IP Statistics

Impact on health:

Lifespan has increased linearly since the 1st Industrial Revolution



Third Industrial Revolution

What are the 3 breakthroughs?

1. Communication:

Computing + Internet / Web

2. Energy:

Solar power

3: Financial:

Crowd funding.

The internet giants become credit banks.

FinTech etc

Third Industrial Revolution

What are the 3 breakthroughs?

1. Communication:

Computing + Internet /

This first component of the TIR is well-ahead of the other two.

2. Energy:

Solar power + New nuclear.

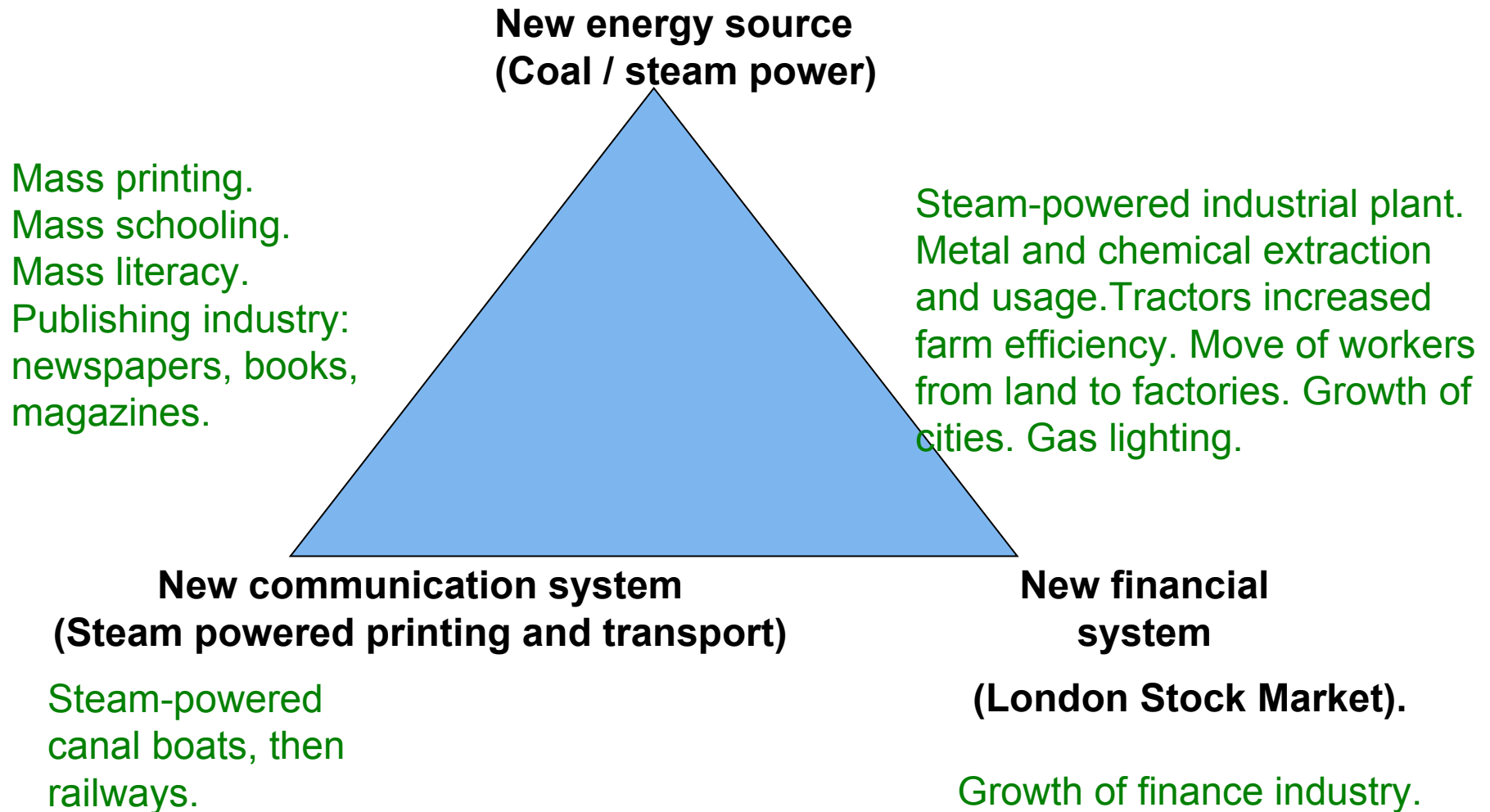
3: Financial:

Crowd funding.

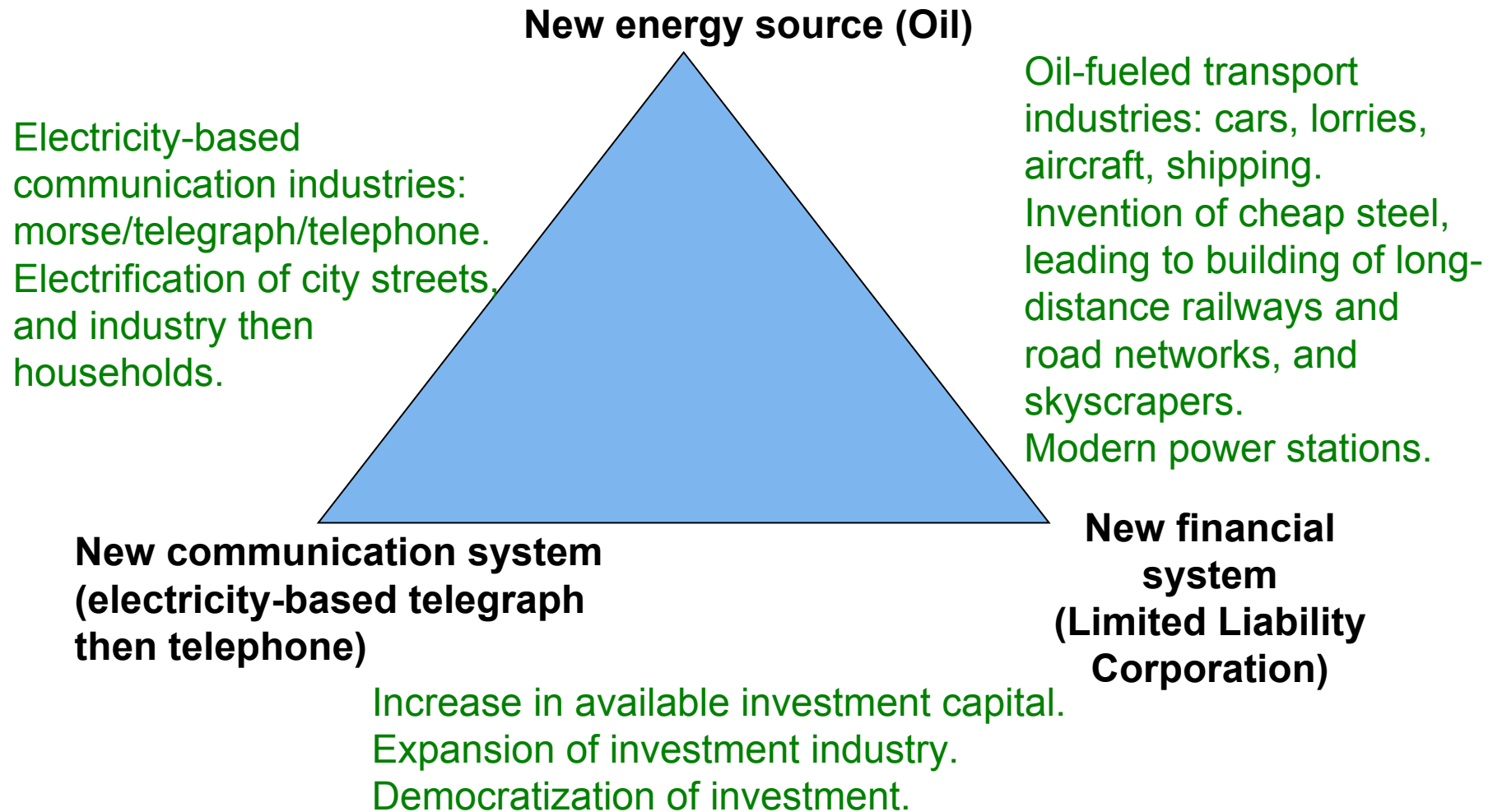
The internet giants become credit banks.

And more to come.

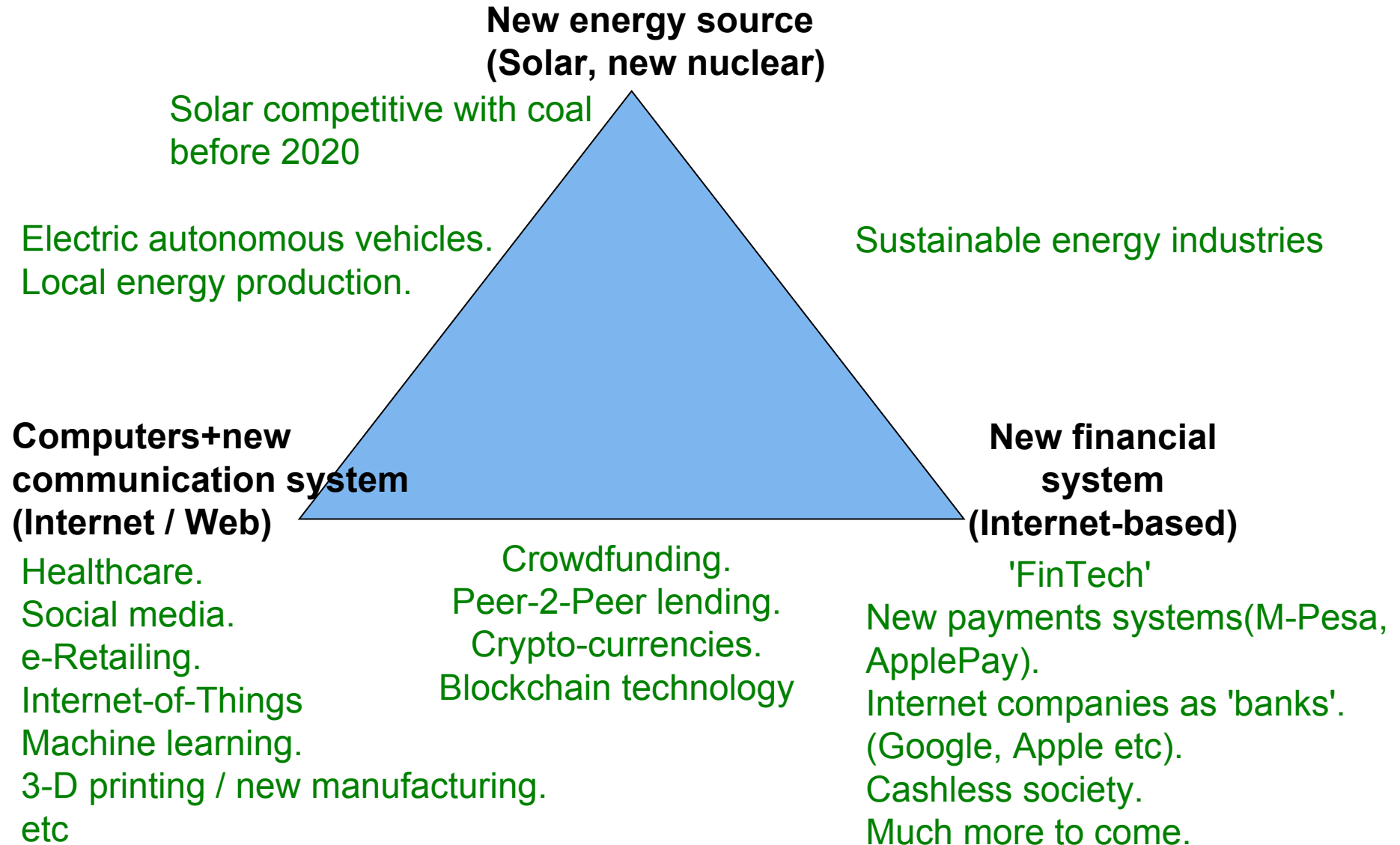
Interaction / synergy between the major drivers of the 1st Industrial Revolution



Interaction / synergy between the major drivers of the 2nd Industrial Revolution



Interaction / synergy between the major drivers of the 3rd Industrial Revolution



Interaction / synergy between the major drivers of the 3rd Industrial Revolution

**New energy source
(Solar, new nuclear)**

Solar competitive with coal
before 2020

Electric autonomous vehicles.
Local energy production.

Sustainable energy industries

**Computers+new
communication system
(Internet / Web)**

Healthcare.
Social media.
e-Retailing.
Internet-of-Things.
Machine learning.
3-D printing / new manufacturing.
etc

Healthcare is the second most advanced area of the TIR, being driven by computing / internet, which is the most advanced of the 3 drivers.

**New financial
system
(Internet-based)**

'FinTech'
New payments systems(M-Pesa, ApplePay).
Internet companies as 'banks'.
(Google, Apple etc).
Cashless society.
Much more to come.

So what is the future scale of big data in healthcare?

Stephens ZD, Lee SY, Faghri F, Campbell, RH, Zhai C, Efron MJ, et al. (2015) Big Data: Astronomical or Genomical?. PLoS Biol 13(7): e1002195. doi:10.1371/journal.pbio.1002195

Compared genomics with three other major generators of Big Data: astronomy, YouTube, and Twitter.

YouTube: 300 hours of video being uploaded every minute
- could grow to 1,000–1,700 hours per minute (1–2 exabytes of video data per year)

Show that genomics is either on par with or the most demanding for data acquisition, storage, distribution, and analysis.

Scale of genomics data

Year 2015

> 2,500 high-throughput instruments, in 1,000 sequencing centers in 55 countries in universities, hospitals and other research laboratories

e.g. Sequence Read Archive (SRA) USA

3.6 petabases of raw sequence data from ~32,000 microbial genomes, ~5,000 plant and animal genomes, and ~250,000 individual human genomes that have been sequenced to date

Genomics sequence data is doubling approximately every seven months

Genomics data storage size could **far exceed** all other domains

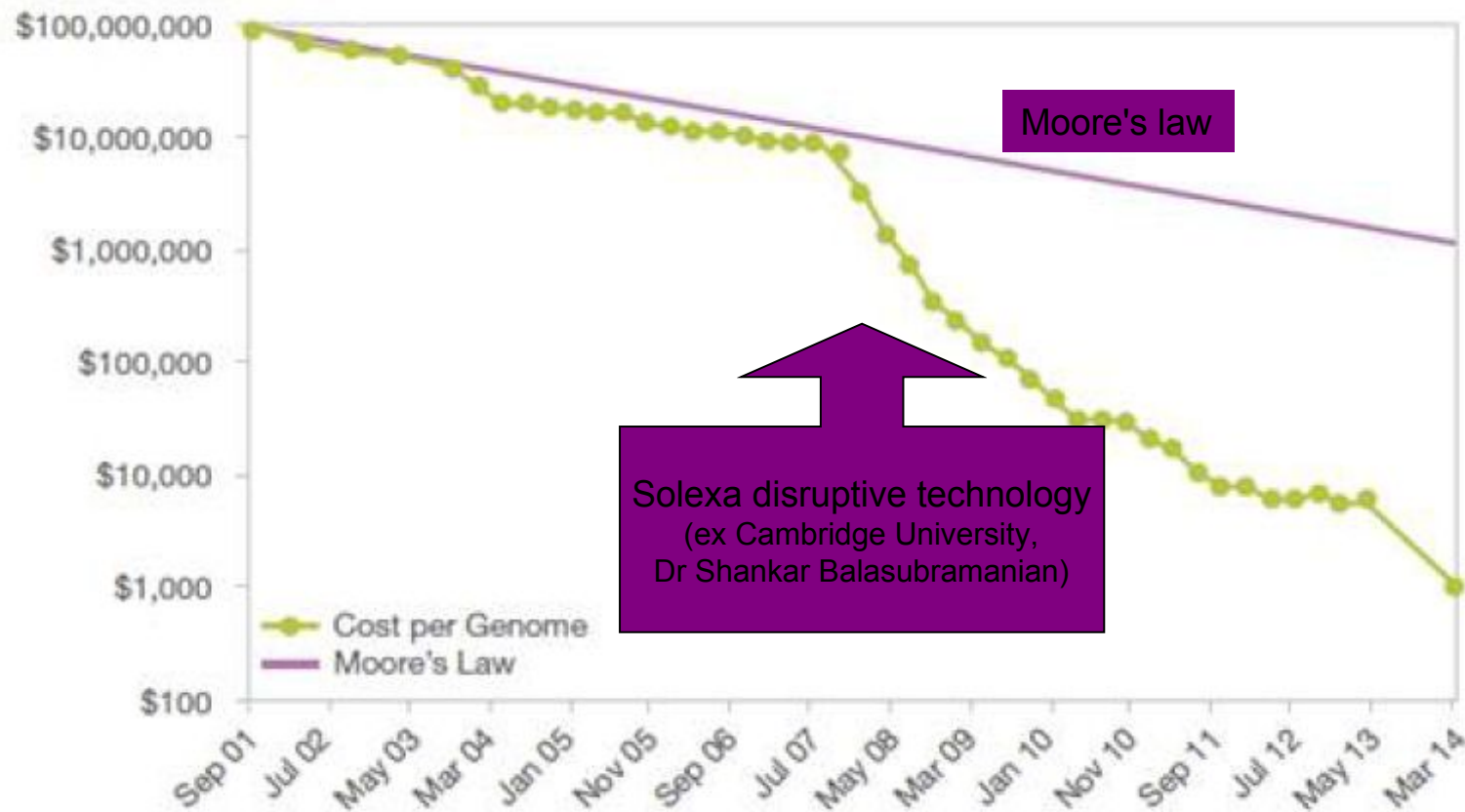
Stephens ZD, Lee SY, Faghri F, Campbell, RH, Zhai C, Efron MJ, et al. (2015)
Big Data: Astronomical or Genomical?
PLoS Biol13(7):e1002195.doi:10.1371/journal.pbio.1002195

We therefore estimate between 100 million and as many as 2 billion human genomes could be sequenced by 2025, representing four to five orders of magnitude growth in ten years and **far exceeding the growth for the three other Big Data domains**. Indeed, this number could grow even larger, especially since new single-cell genome sequencing technologies are starting to reveal previously unimagined levels of variation, especially in cancers, necessitating sequencing the genomes of thousands of separate cells in a single tumor [10].

Faster than Moore's Law

"Gene-sequencing technology is advancing at a rate even faster than computer processing power."

McKinsey 2013



Personal genome sequencing now below \$1k

\$2.7 billion, 13 years

Cost and duration of the Human Genome Project,
completed in 2003

Current cost \$1000 per genome.

2,700,00 cost reduction in 12 years.

225,000x improvement per year.

\$100, 1 hour

Cost and time to sequence a human genome
in the next decade²

Growth of DNA sequencing

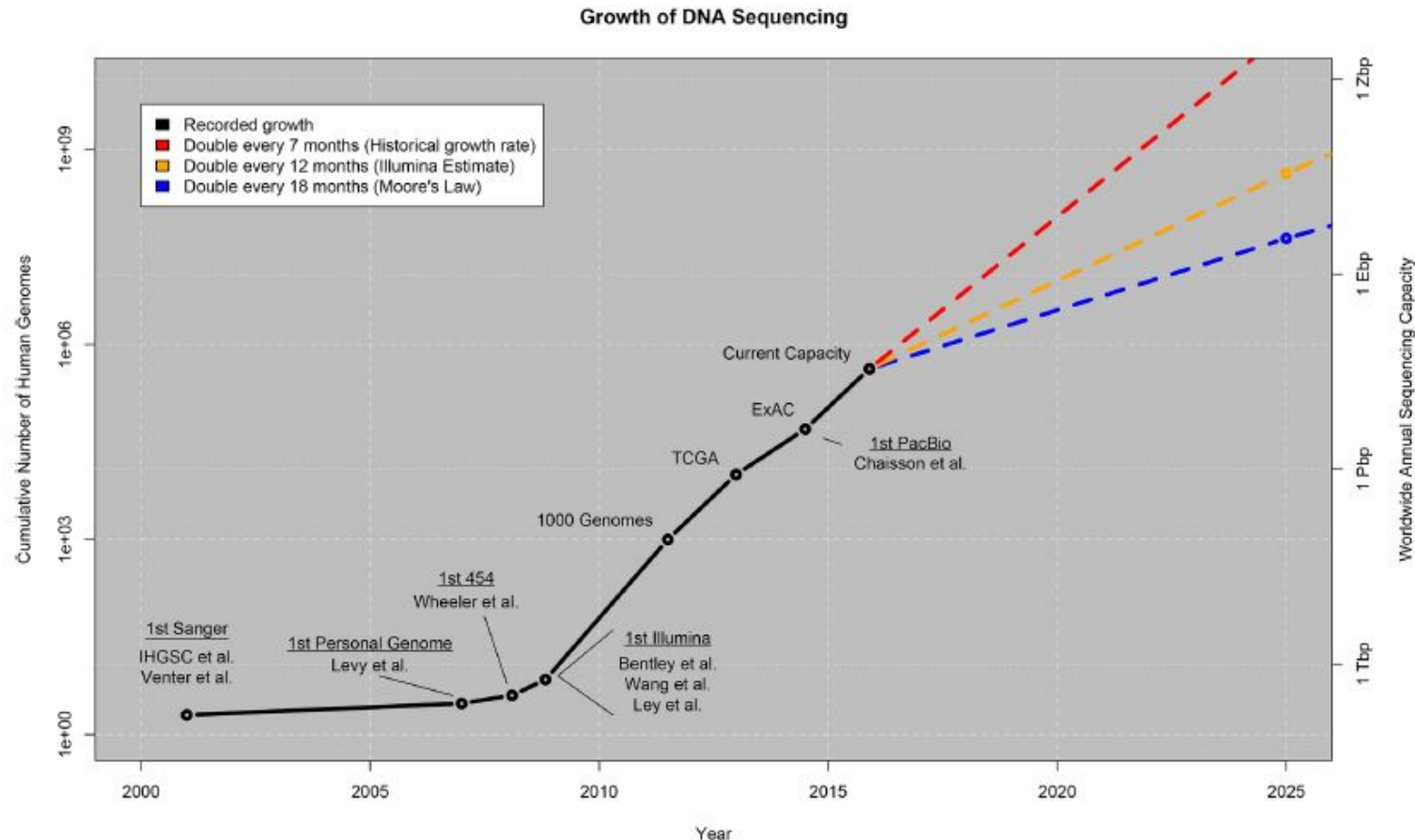


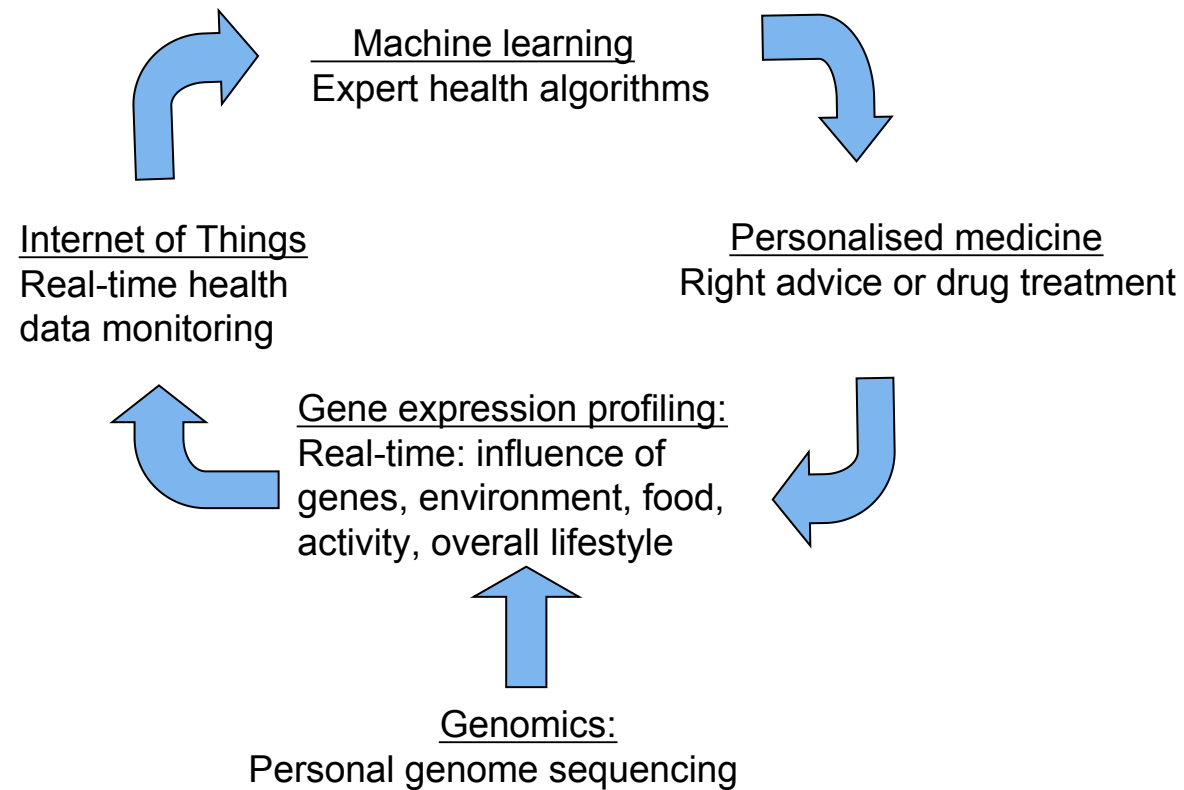
Fig 1. Growth of DNA sequencing. The plot shows the growth of DNA sequencing both in the total number of human genomes sequenced (left axis) as well as the worldwide annual sequencing capacity (right axis: Tera-basepairs (Tbp), Peta-basepairs (Pbp), Exa-basepairs (Ebp), Zetta-basepairs (Zbps)). The values through 2015 are based on the historical publication record, with selected milestones in sequencing (first Sanger through first PacBio human genome published) as well as three exemplar projects using large-scale sequencing: the 1000 Genomes Project, aggregating hundreds of human genomes by 2012 [3]; The Cancer Genome Atlas (TCGA), aggregating over several thousand tumor/normal genome pairs [4]; and the Exome Aggregation Consortium (ExAC), aggregating over 60,000 human exomes [5]. Many of the genomes sequenced to date have been whole exome rather than whole genome, but we expect the ratio to be increasingly favored towards whole genome in the future. The values beyond 2015 represent our projection under three possible growth curves as described in the main text.

The figures on previous slides may be a massive under-estimate. Why?

- Genomes sequenced to date have been whole exome. **Future focus will be whole genome.**
- More than genome will be needed, much more!
Phenotype data also required. Gene expression data, physiological changes, real time diagnostics and prognostics, feedback etc
- And eventually **live data collection and analysis** will be in demand for most humans via Internet-of-Things
 - Data from embedded systems, the signals from which are a major component of the Internet of Things, will grow from 2% of the digital universe in 2013 to 10% in 2020. Healthcare will push it much higher in the 2020s.

The future will be driven by...

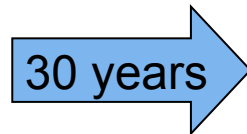
The health learning circle



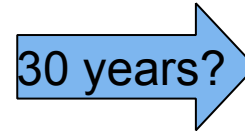
Do you doubt this is possible?



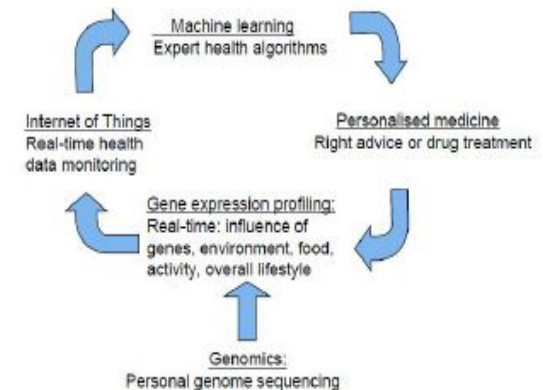
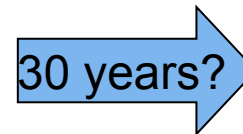
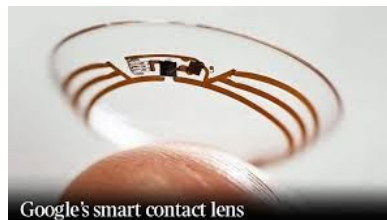
1984:
64k memory.
Home hobbyist



More computing
power than whole
USA 50 yrs ago



More computing power
than whole world today



I-o-T in Healthcrae

- Fed by sensors (wearables, clothing etc)
 - Billions likely eventually for healthcare alone
- Millions of 'intelligent apps'
- Real time data collection and analysis
- Massive data stotage and analysis requirements
- Will drive new consumer goods and business opportunities
- Trillion dollar market
- Reorganisation of healthcare industry

Progress requires the following

- Information security
- Seamless public and private cloud computing
- Agreement on standards
- Next- generation analytics
- New storage management technologies
- New data access tools and processes
- Automatic tagging
- Ability to deal with real-time data

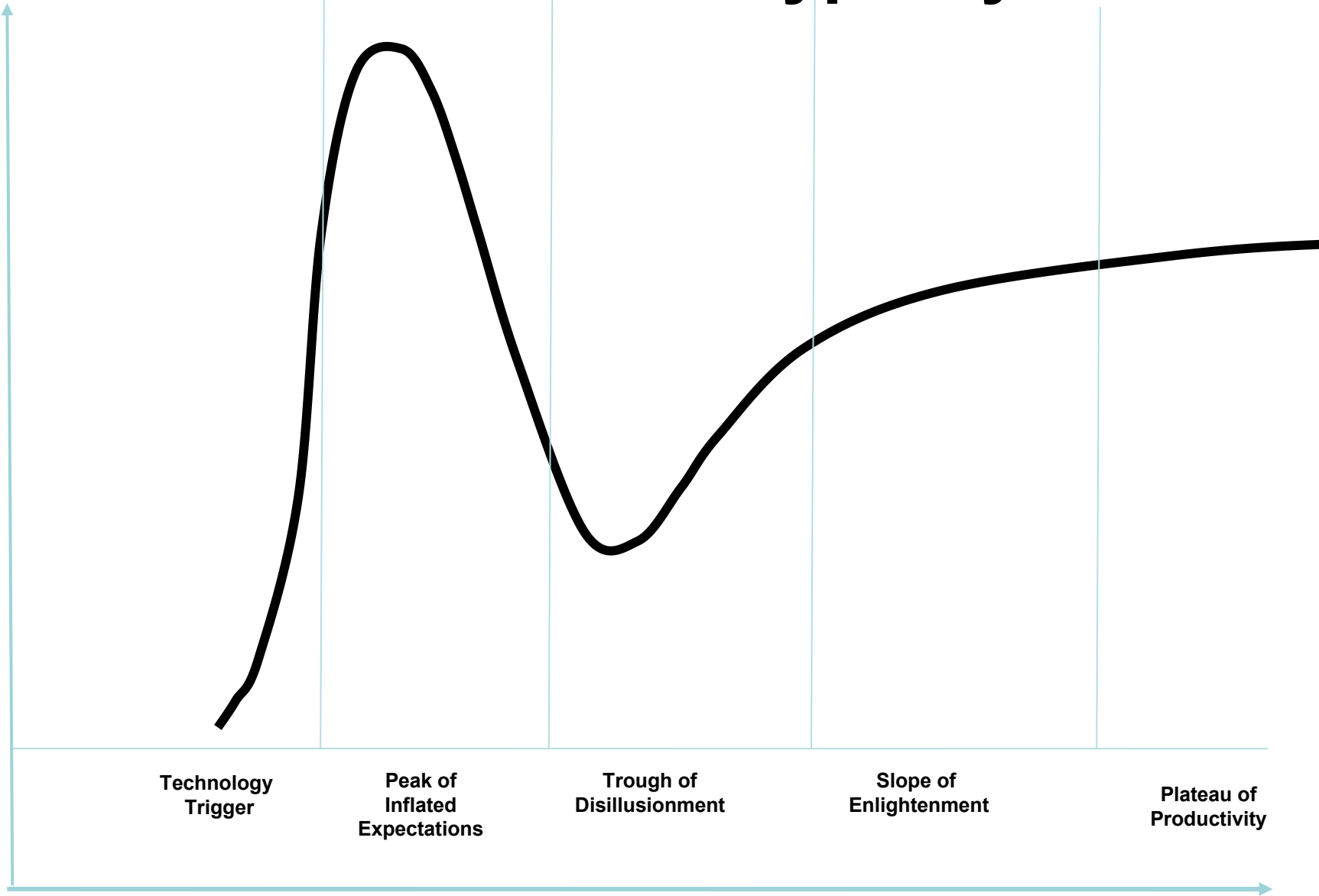
But when?

Where are we now?

How fast will healthcare change?

The Gartner Hype Cycle

Expectations



Technology
Trigger

Peak of
Inflated
Expectations

Trough of
Disillusionment

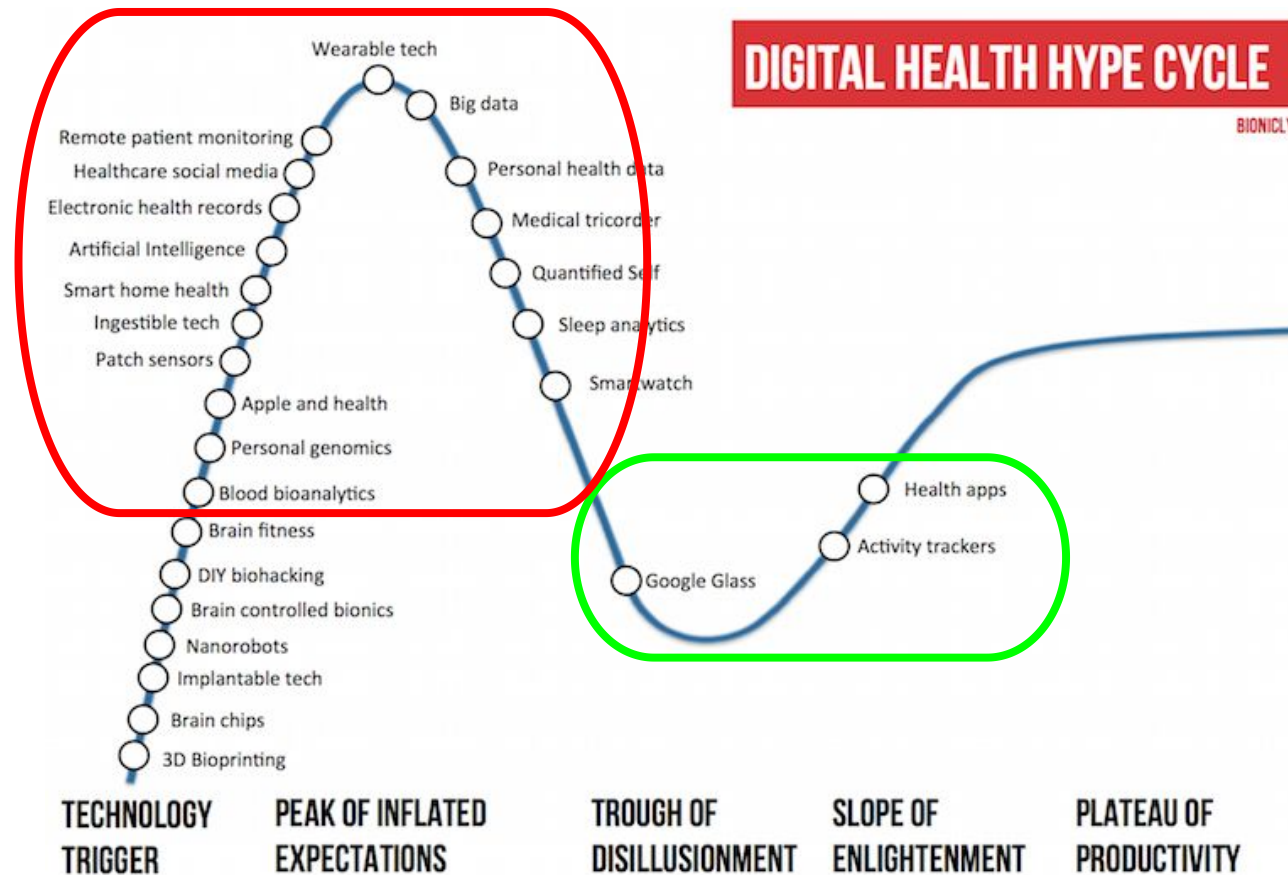
Slope of
Enlightenment

Plateau of
Productivity

Time

Stephen Davies and colleagues applied the Gartner hype cycle.

(Published December 2014)

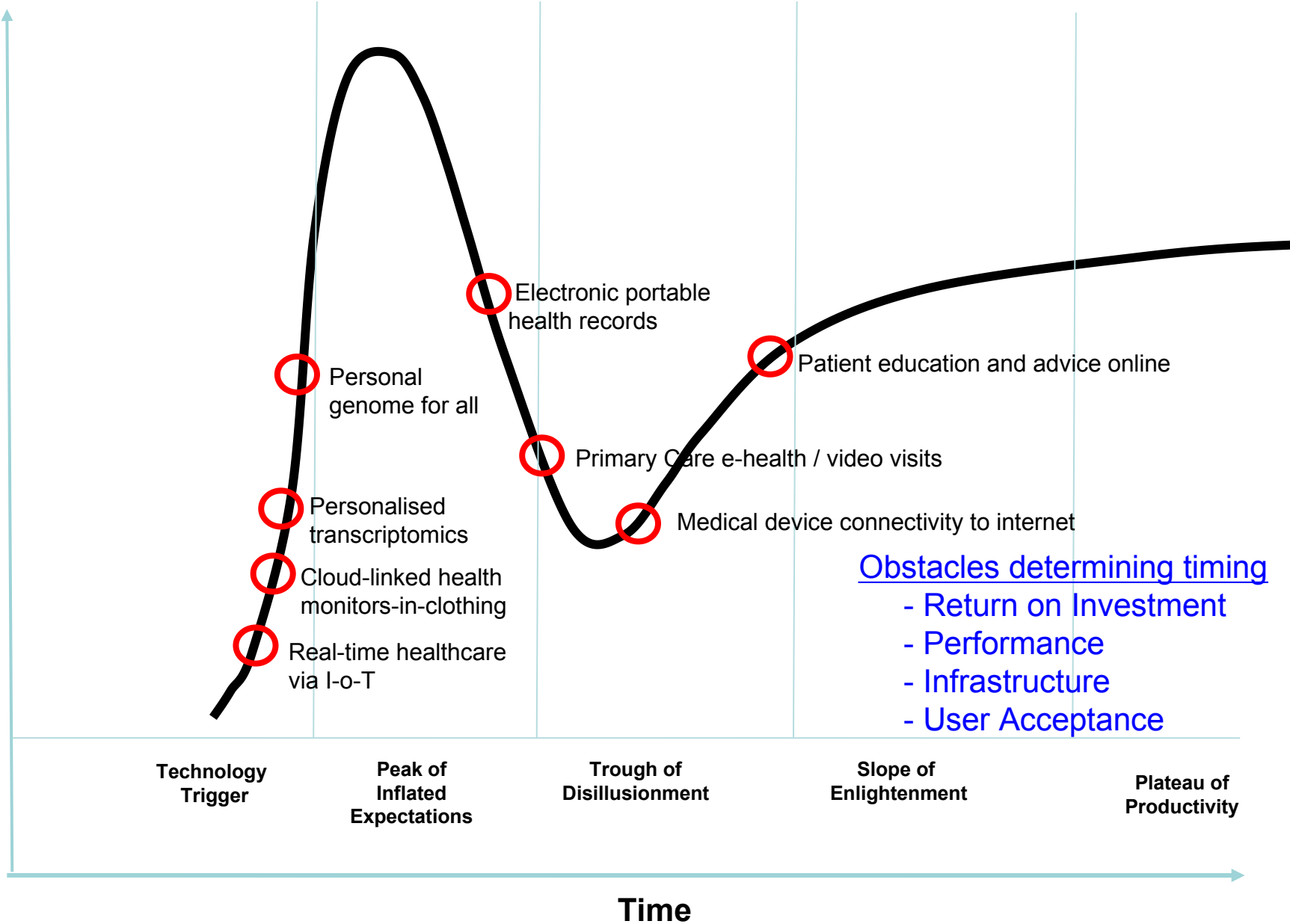


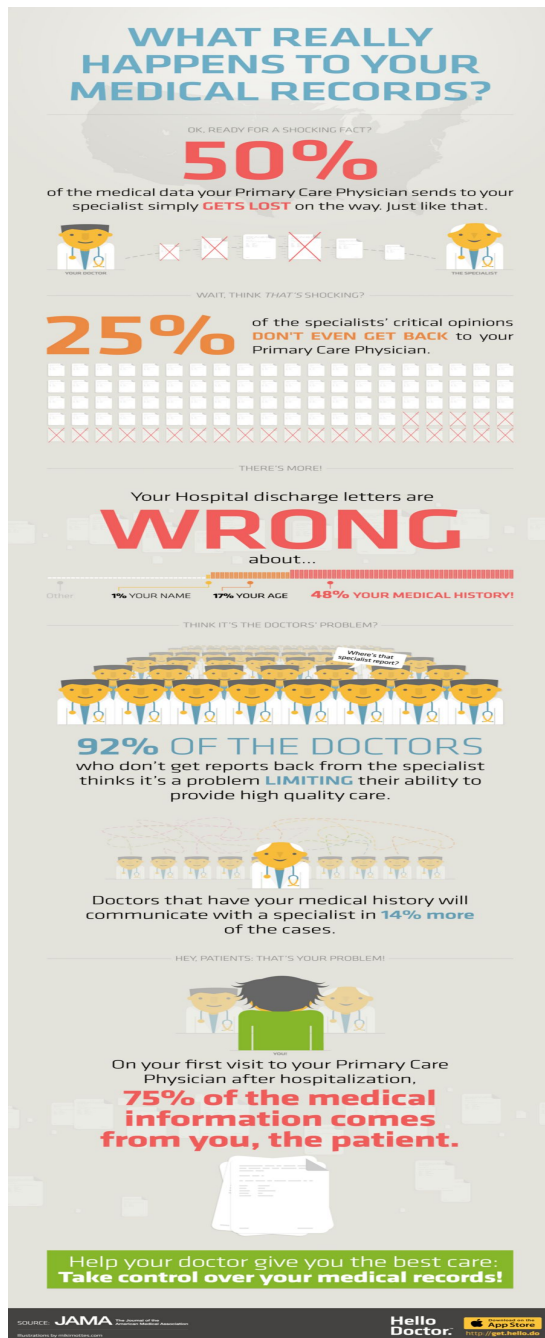
This Digital Health Hype Cycle has in no way been endorsed by Gartner, Inc.

Reference: <http://bionicly.com/digital-health-hype-cycle/>
NB. The digital health hype cycle has not been approved by Gartner!

Expectations

And I would add...





We really do need change...

This is where we are now.

JAMA on your medical record

Primary care → Specialist:
50% of data from primary care to specialist gets lost.

Specialist → Primary care:
25% specialist opinions do not get back to primary care.

Hospital → Primary care
48% hospital discharge records are wrong on your medical history
75% of your medical information after hospitalisation comes from you personally on next visit primary care visit.

The Infographic was created by Hello Doctor and is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License and shared here with Hello Doctor's full permission.

Source data from the Journal of American Medical Association (JAMA) in these two studies:
<http://jama.jamanetwork.com/article.aspx?articleid=205790> and
<http://archinte.jamanetwork.com/article.aspx?articleid=226367>

The winners in digital healthcare may not be
the current dominant healthcare players

Business Model Innovation is
More Powerful than Technology Innovation.

Those Without a Legacy Have
Fewer Barriers to Adopting Radical Innovations

Three business models will merge:

- Pharma/biotechnology
- Diagnostics
- Internet

...creating a totally new business model

3rd International Systems Biomedicine Symposium
“Big Data in Health Care”
28-29 October, 2015, Luxembourg

Big Data in Healthcare: What impact over the next decade?

Dr David Brown, PhD, FRSC, FRSM

Alchemy Biomedical Consulting
Cambridge, UK
alchemybiomedicalconsulting.vpweb.co.uk
email: davidbrown1000@btinternet.com
+44 (0) 7766 686 345