

How I Learned to Stop Worrying and Love the Imminent Internet Singularity

Or: Why right now may be the best possible time to be an
Internet technologist

Gary William Flake
Microsoft / MSN

Outline for Today

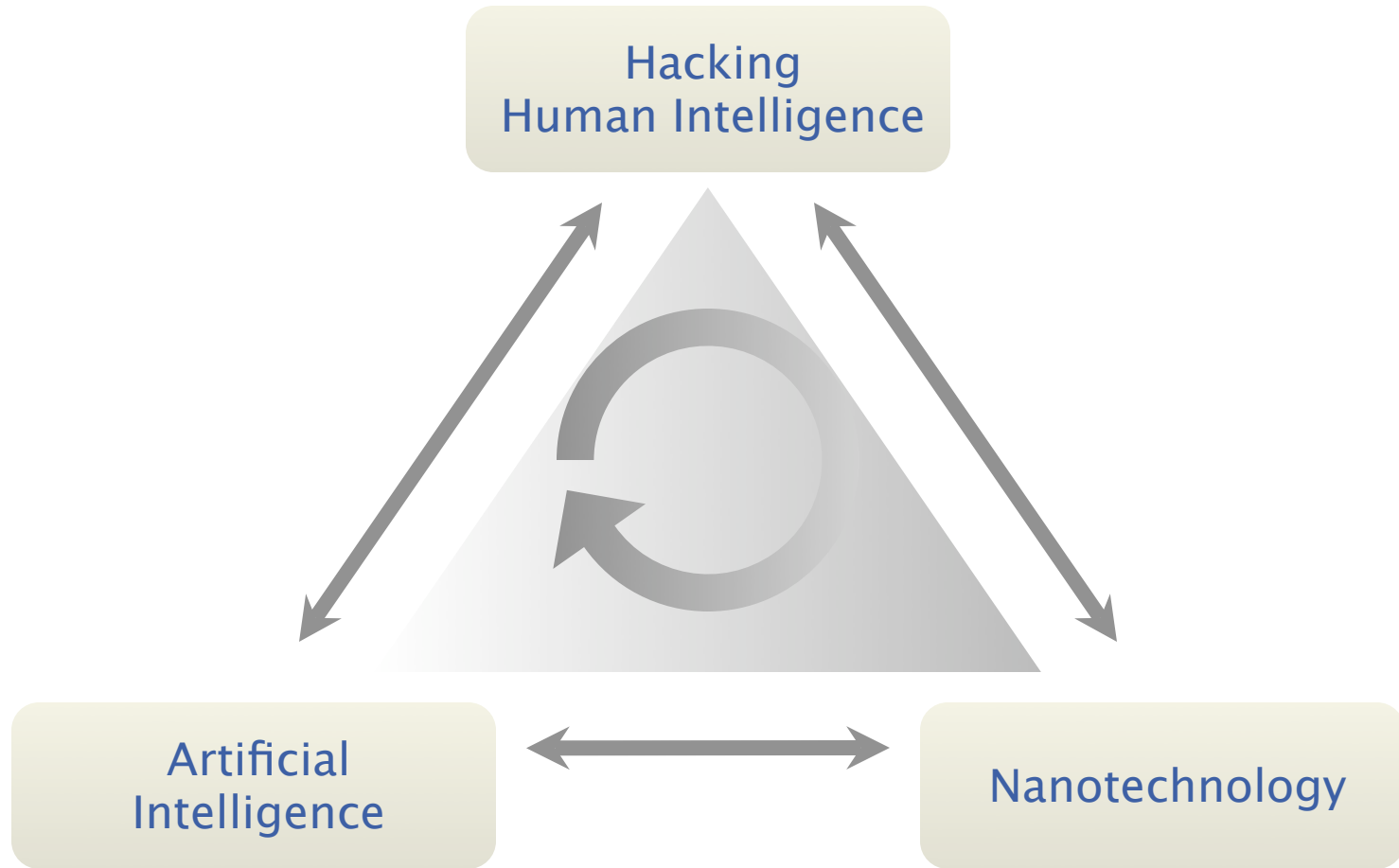
- **Internet Singularity**
 - Why right now is the best time in the history of the universe to be an Internet technologist.
- **Announcing Live Labs**
 - A partnership between MSN and MSR

Back Story

Singularity: the idea that the exponential speed of technological improvements will produce super-human capabilities, making the future completely unknowable.

- Introduced by Vernor Vinge, in 1993.
- Evangelized by Ray Kurzweil in multiple books.
- Name purposely suggests black hole event horizon.
- Primary causes cited as the realization of true AI and nanotechnology, yielding asymptotic upper limits of intelligence, power, and ability to manipulate reality.
- Corollary: human society will end as we know it.
- Talked about in both optimistic and pessimistic

Multiple Paths to the Singularity



The Internet Singularity

Loosely defined: the idea that a deeper and tighter coupling between the online and offline worlds will accelerate science, business, society, and self-actualization.

- Coined by me. Also called the “Baby Singularity”.
- Primary cause is claimed to be ubiquitous computing, democratization of computing resources, and iterative processes of creation and discovery becoming continuous.
- Corollary: the increased pace of discovery will fuel amazing innovations in the near future.
- Will completely reshape society as we know it.
- A kinder and gentler singularity. Very optimistic statement.

Making the Case in Four Steps

1. Democratization and “macro-ization”
2. Power law distributions and “long tails”
3. Internet ecosystems and network effects
4. The Innovators’ Dilemma

(1/4) Democratization & “Macro-ization”

- Personal & business computing
 - Greater availability of raw CPU power
 - Today’s desktop > Yesterday’s super computer
 - Tomorrow’s cell phone > Today’s desktop
 - Greater availability of powerful software tools
 - Greater integration among toolsets
- Largest changes occurring in
 - Content
 - Commerce
 - Community
- Positive feedback loop from virtuous

(1/4) Democratization of Content

- Documents: Office, LaTeX, HTML
- Images: digital cameras, Photoshop, iPhoto
- Movies: camcorders, DVDs, iMovie, Xbox
- Audio: MP3, GarageBand, podcasting
- Publishing: blogs, RSS
- Software: open source, VB, Flash
- Research: search engines, wikipedia
- Meta Data: tags, play lists, recommendations

(1/4) Democratization of Commerce

- Hosting: Y! Stores, Office Live, Amazon
- P2P Sales: eBay
- Transactions: PayPal
- Marketing: Overture, Adwords, AdCenter
- Syndication: Adsense, RSS
- Development: Web APIs
- Trivial Work: Mechanical Turk, ESP game
- Business Intelligence: Web analytics

(1/4) Democratization of Community

- Communication: IM, email, VOIP
- Distribution: RSS, BitTorrent, P2P
- Social: Friendster, blogs, Y/G/M groups, Meetup
- Romance: Match.com, eHarmony
- Virtual Economies: Everquest, WoW
- Reputation: eBay, PageRank
- Preference: collaborative filtering
- Affinity: implicit relationships via similarity

(1/4) Democratization & “Macro-ization”

Major Themes:

- Today’s “amateur” has more resources than yesterday’s “professional”.
- The difference between “amateurs” and “professionals” is diminishing over time.
- The number of producers / creators is increasing in both absolute and relative terms.

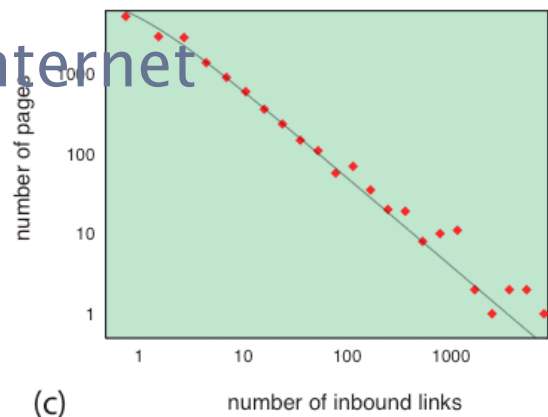
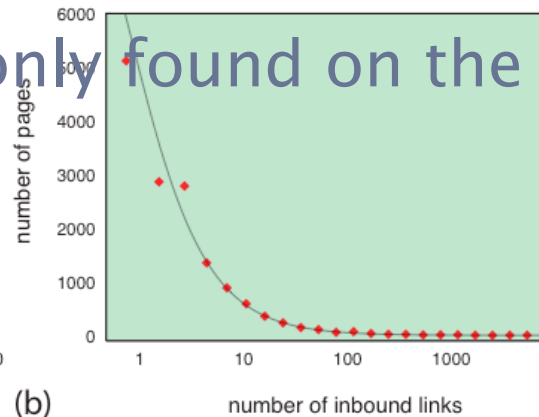
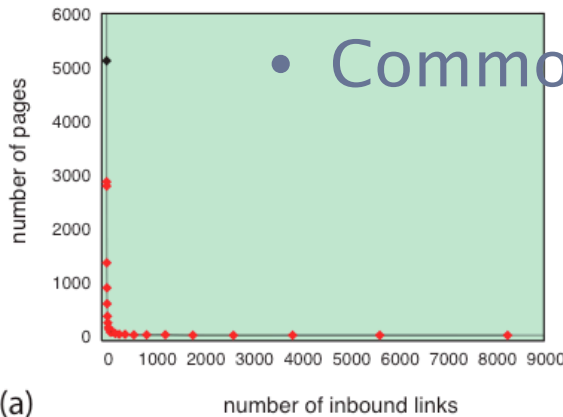
Implications:

- The distribution of “creators” is dramatically growing and radically changing in structure.

(2/4) Power Laws & “Long Tails”

Power Laws:

- Exponentially more small things than large things.
- Canonical example: distribution of biomass
 - Small number of whales and elephants
 - Enormous number of bacteria



• Commonly found on the Internet

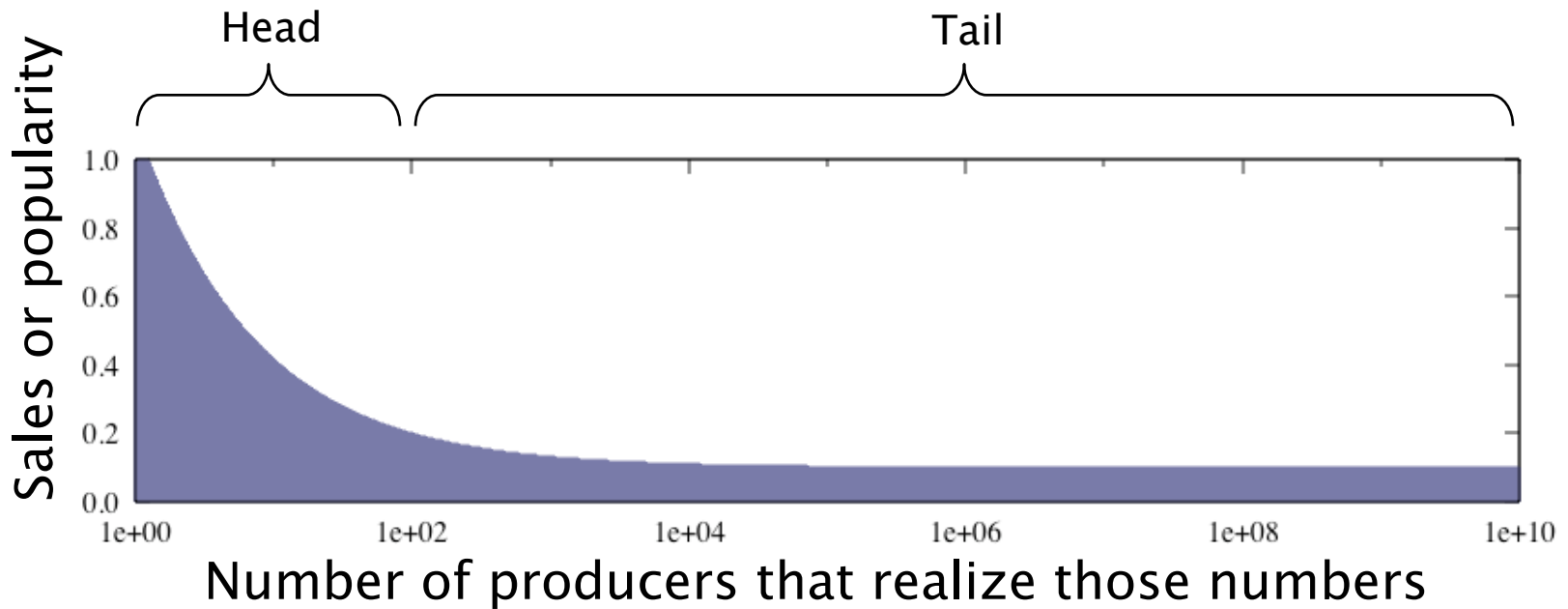
(2/4) Power Laws & “Long Tails”

Long Tails:

- The “weight” of the tail is greater than the head.
- Coined by Chris Anderson in Wired article of the same title.
- Canonical example is the recording industry:
 - Few “Britney Spears” (i.e., performers with enormous record sales).
 - Vast number of unsigned performers with little or no record sales.
- Requires that “tail” producers be able to survive.

(2/4) Power Laws & “Long Tails”

Visualizing Long Tails:



(2 / 4) Power Laws & “Long Tails”

Major Themes:

- Long tails found in content, commerce, and communities.
- Exist because “physics” of the online world differs from the offline world:
 - Size of warehouse and shelf space is irrelevant.
 - Distance and medium are irrelevant.
 - Aggregation, remixing, and tagging all contribute.
 - Scale can increase independently of human action.

Implications:

- Consumers are slowly becoming producers.
- In combination, the small producers may outweigh the large producers.

(3 / 4) Ecosystems & Network Effects

Ecosystems – participants on all scales

- Head participants – old school companies
- Tail participant – “prosumers”
- Aggregations and playlists
- Remixes and mashups
- Annotations, reviews
- Tags and meta data
- Mere activity

(3 / 4) Ecosystems & Network Effects

- **Network Effect**
 - The phenomenon whereby the value of a network increases as a function of the number of participants.
- **Direct Network Effect: Telephones**
 - The more people that have one, the more valuable each phone is to each user.
- **Indirect Network Effect: OS Development**
 - Users use an OS because it has the most applications.
 - Developers write to an OS because it has the most users.

(3 / 4) Ecosystems & Network Effects

- Major Theme:
 - The whole > sum of the parts
 - Content: the language of preference.
 - Commerce: promotes information flow through endorsements, reputation and market signals.
 - Community: acts as a collective filtering mechanism.
- Implications:
 - Each is the “chicken” to the other’s “egg” in that lacking any makes the other two less valuable.
 - In combination, networks mutually reinforce one another.

(4/4) The Innovators' Dilemma

The Pattern:

- First in industry focus on small number of large and high-margin customers.
- Late arrivals have to focus on lower-margin customers.
- Late arrivals learn efficiencies because they compensate lack of margin with scale.
- Meanwhile, competition increases, margins shrink.
- The established companies rarely learn the efficiencies that the younger companies grew up with.
- The late arrivals win because they can take optimizations and apply them to the head.

(4/4) The Innovators' Dilemma

- The Dilemma Defined:
 - The first companies in an industry (the innovators) must eventually destroy their own business before someone else does.
- Why? Disruptions happen from the bottom:
 - Cray killed by SGI; SGI killed by Sun; Sun killed by PC; (Will the PC be killed by cell phones?)

(4/4) The Innovators' Dilemma

Offline World

- Huge startup costs for new businesses.
- Aggregate size of the tail is limited by physics.
- More business usually implies more employees (work harder).
- Quality product usually implies high touch.
- Innovation iterations follow product and business cycles.

Online World

- Diminishing startup costs for new businesses.
- Aggregate size of tail potentially unlimited.
- More business may not require more employees (work smarter).
- Quality product can be a better algorithm.
- Innovation iterations follow data flow cycles.

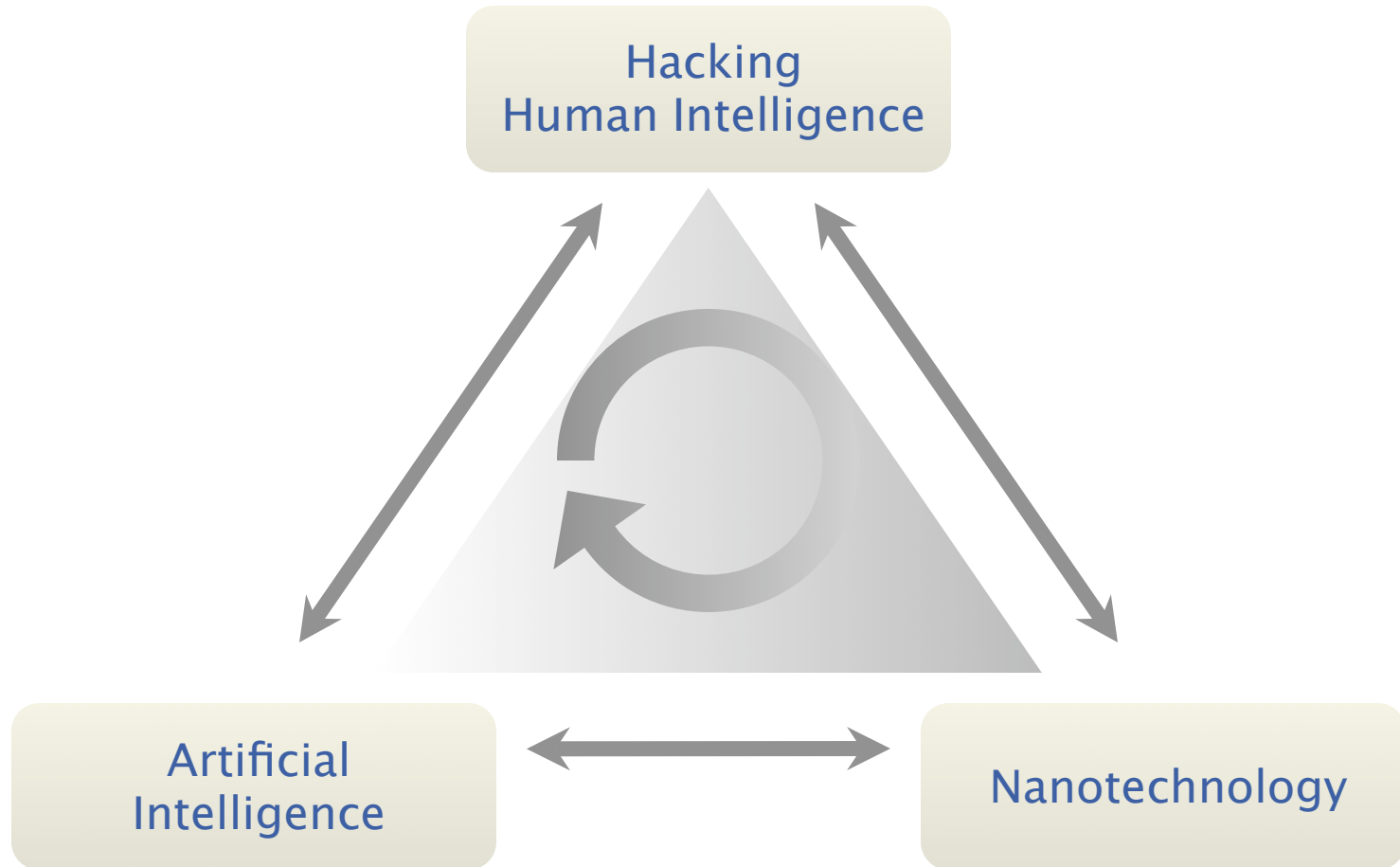
(4/4) The Innovators' Dilemma

- Major Themes:
 - A bigger tail makes for more potential disruptions.
 - The non-physical aspects of the Internet speed up the natural “clock cycle” of progress.
- Implications:
 - The process of societal evolution is itself changing.

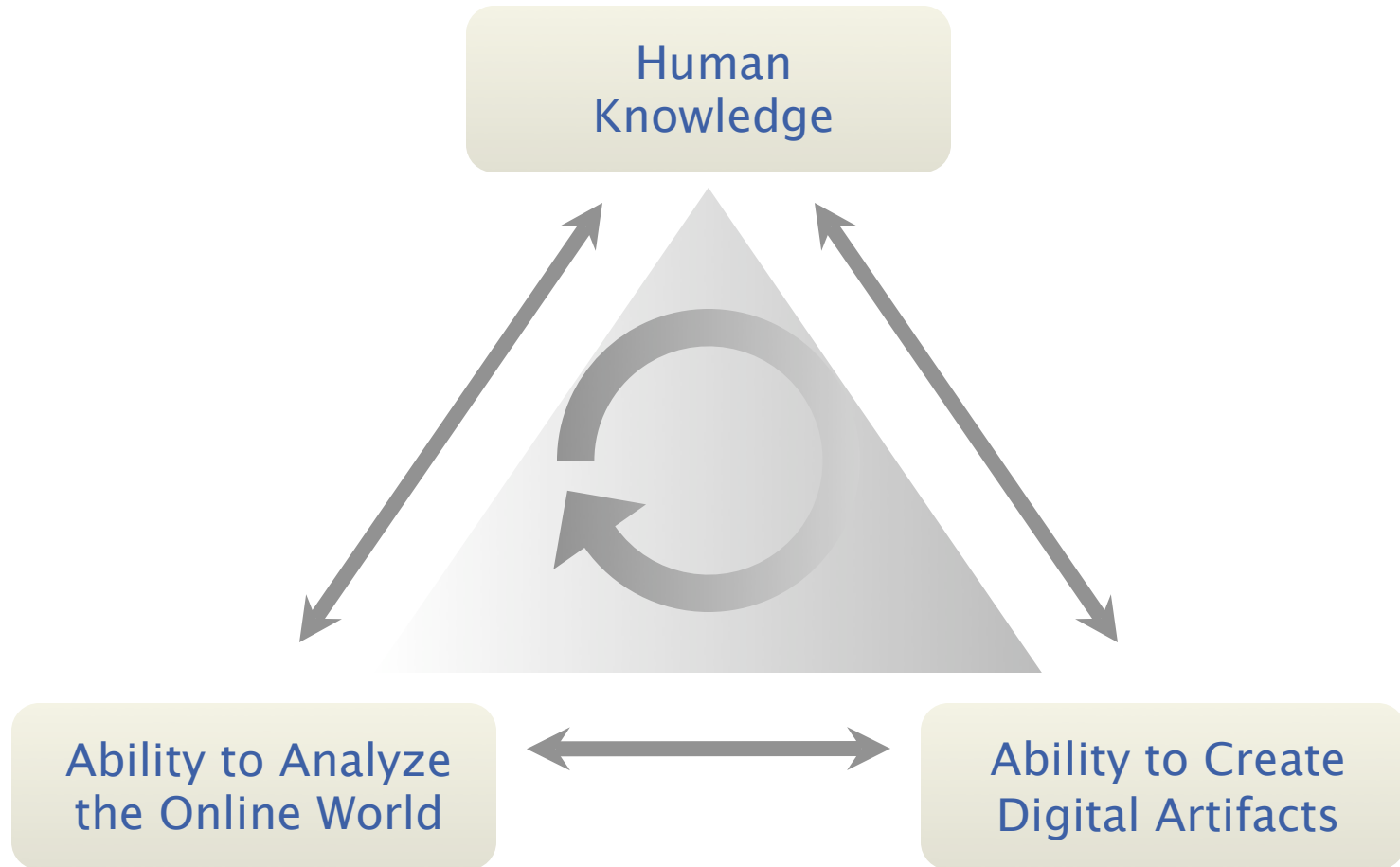
Let's Recap

1. Democratization and “macro-ization”
 - Yields massive parallelization.
2. Power law distributions and “long tails”
 - Decentralization is bigger than centralization.
3. Internet ecosystems and network effects
 - The whole > sum of the parts.
4. The Innovators' Dilemma
 - Higher bandwidth and lower latencies on information flow.
 - More frequent improvements.

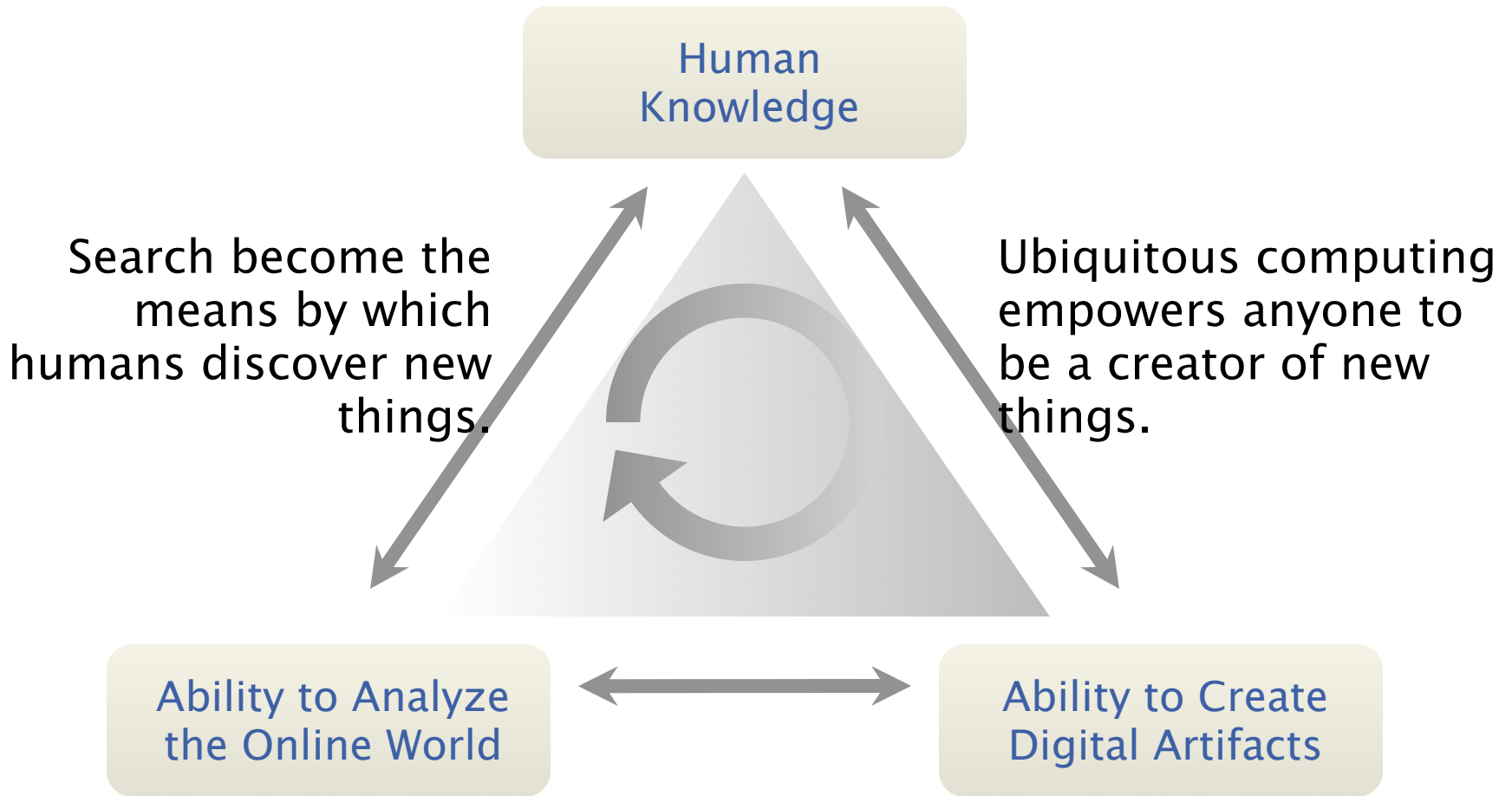
Remember This Slide?



Let's Change the Labels



The Internet Singularity



The relationships between objects will inform our understanding of the world.

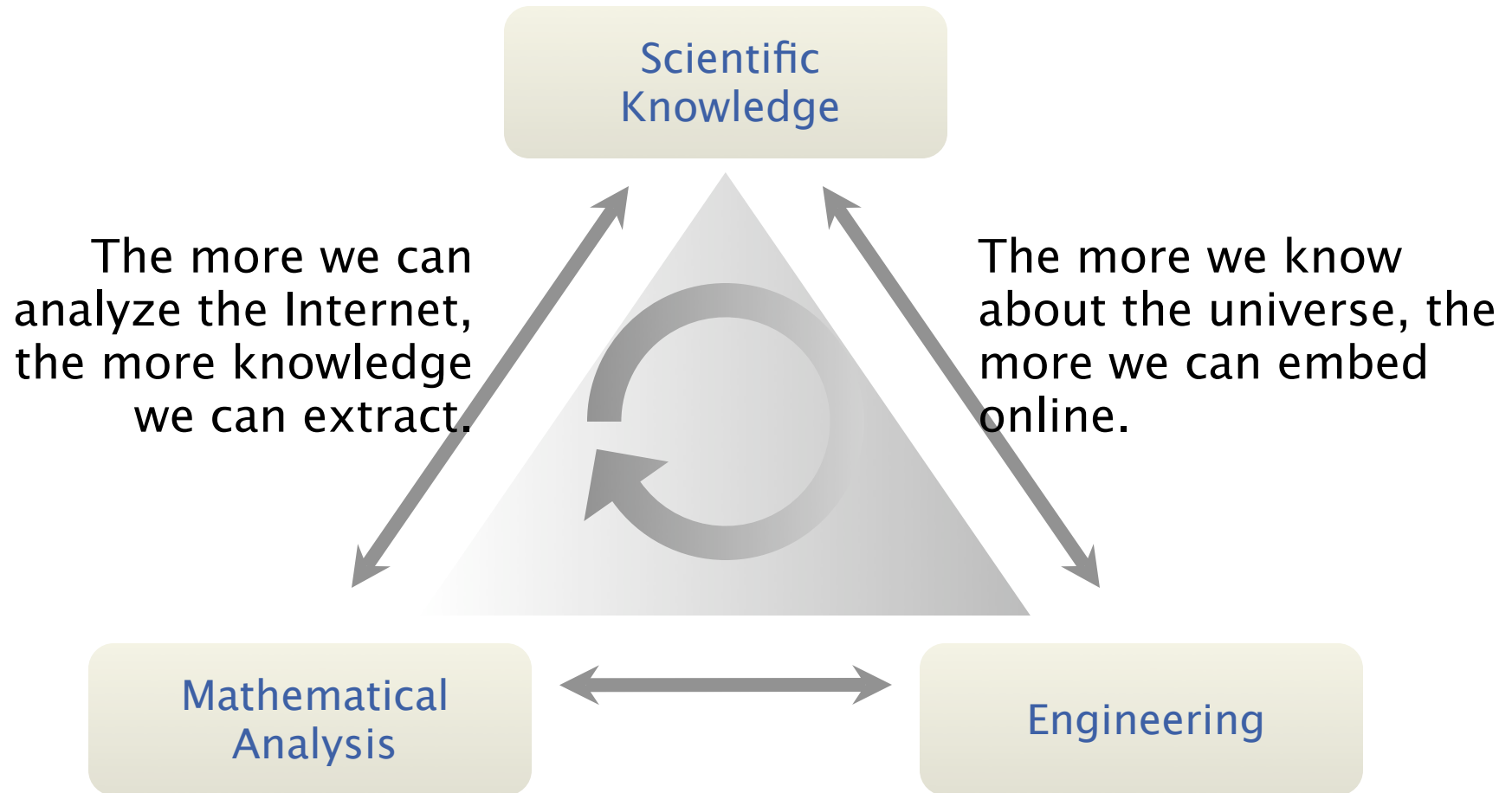
The Internet as a Mirror

- As time goes on, the Internet's content, composition, and participants more accurately reflects the physical world.
- In the limit, the physical world effectively becomes instrumented by the virtual world.
- Having this strong connection between the online and offline worlds allows for science to be carried out in a revolutionary new way:
 - Theories will be tested relative to Internet data.
 - Empiricists will have almost unlimited data.
 - Simulations will truly allow us to experiment in a universe of theories.

Existing Evidence

- Web search is now the greatest applied CS R&D problem and it will drive the priorities of AI and other sub-fields for decades to come. The future of the search engine is to model the human mind in the aggregate.
- Large corpora are revolutionizing NLP.
- Virtual worlds are now used to study emerging economies.
- The online bestiary (viruses, Trojan horses, etc.) is now a legitimate way to study epidemiology.
- Empirical sociology is now carried out on Web data.
- NIH is now planning on the day for when medical literature can be mined for medical

The Computer Science Singularity



The online world lends itself to more thorough analysis.

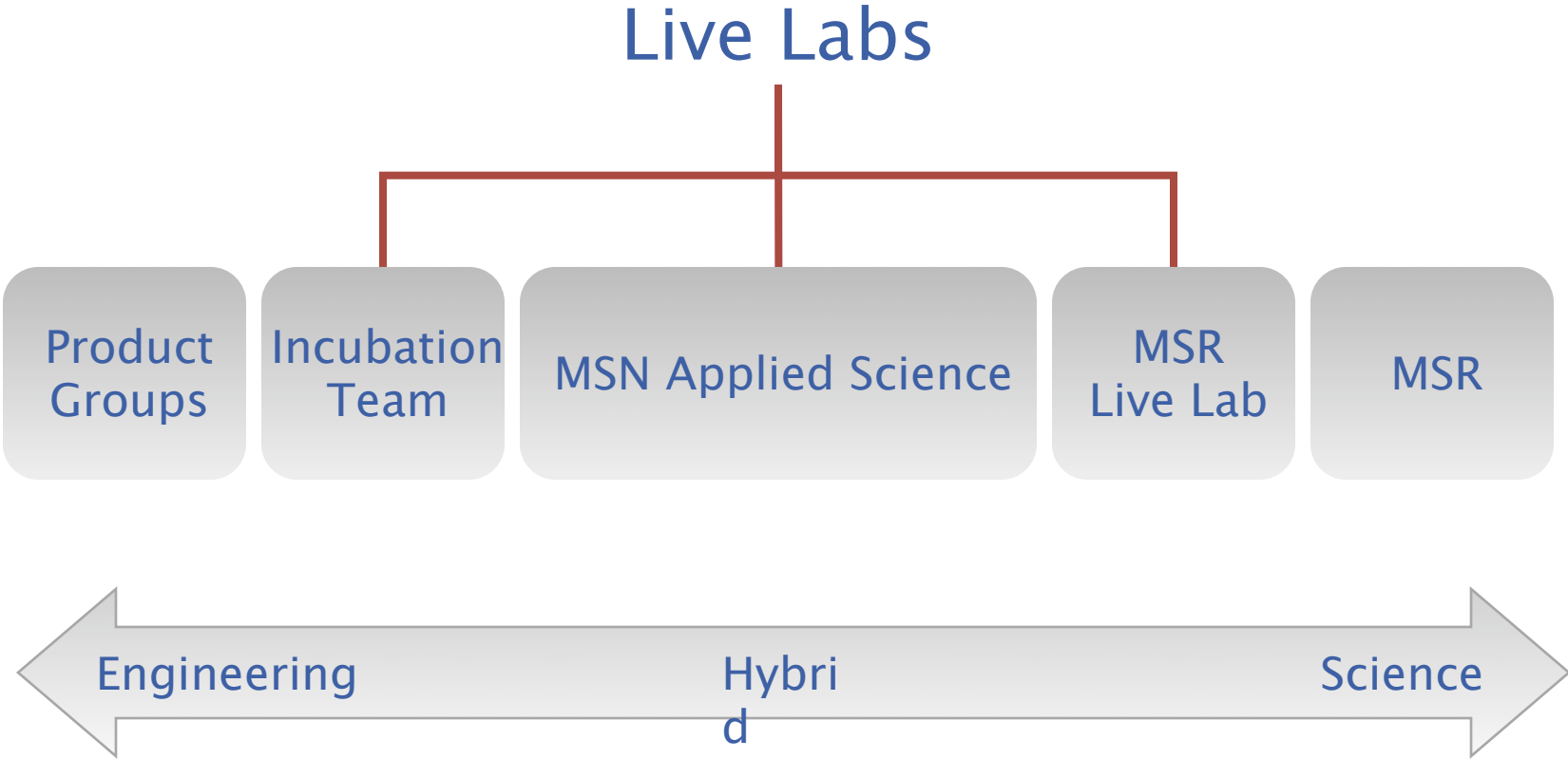
Outline for Today

- Internet Singularity
 - Why right now is the best time in the history of the universe to be an Internet technologist.
- **Announcing Live Labs**
 - A partnership between MSN and MSR

Live Labs

- Partnership between MSR and MSN.
- 100+ new positions but also a “virtual” organization
- Aspires to hit the “sweet spot” between:
 - Science and engineering
 - Top-down and bottom-up
 - Users and businesses
 - Short-term and Long-term
 - Problems and solutions
- Incubates ideas, projects, and people.

Approximate Structure



Live Labs – High Level Goals

- Create a virtuous cycle in three parts:
 - Build great products
 - Rapidly expose products to the outside world
 - Have a high bandwidth channel from the outside world back in
 - Repeat
- We want this process to be continuous (instead of discrete).

Live Labs – Secondary Goals

- Hire more talent!!!
- Invest in more partnerships
- Make more engineering resources available to MSR.
- Create an applied science career track within MSN.
- Improve our innovation clock cycle.
- Allow for efforts outside of the product roadmap.
- Reserve resources for rapid prototyping.
- Reserve resources for R&D platforms.
- Bottom line: to fully make use of how the

Final Remarks

- We have a vision for how the Internet is ushering in profound changes that will benefit everyone.
 - This vision suggests major changes to all of society.
- Our mission is to create and nurture the virtuous cycle from creation → publication → feedback.
- Our strategy is to create an institute that hits the “sweet spot” on many dimensions to more effectively bridge between worlds.