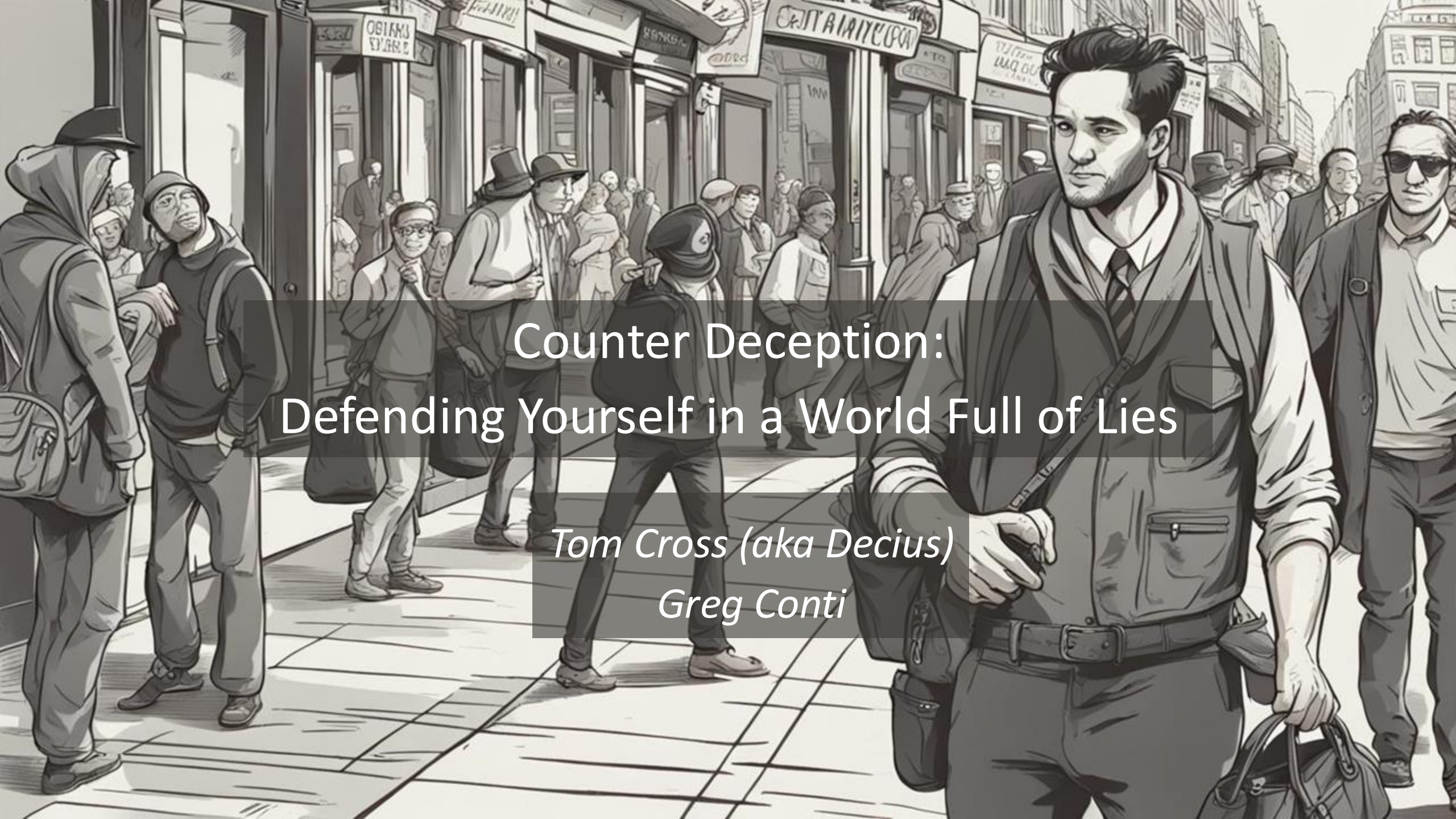


Counter Deception:
Defending Yourself in a World Full of Lies

Tom Cross (aka Decius)
Greg Conti



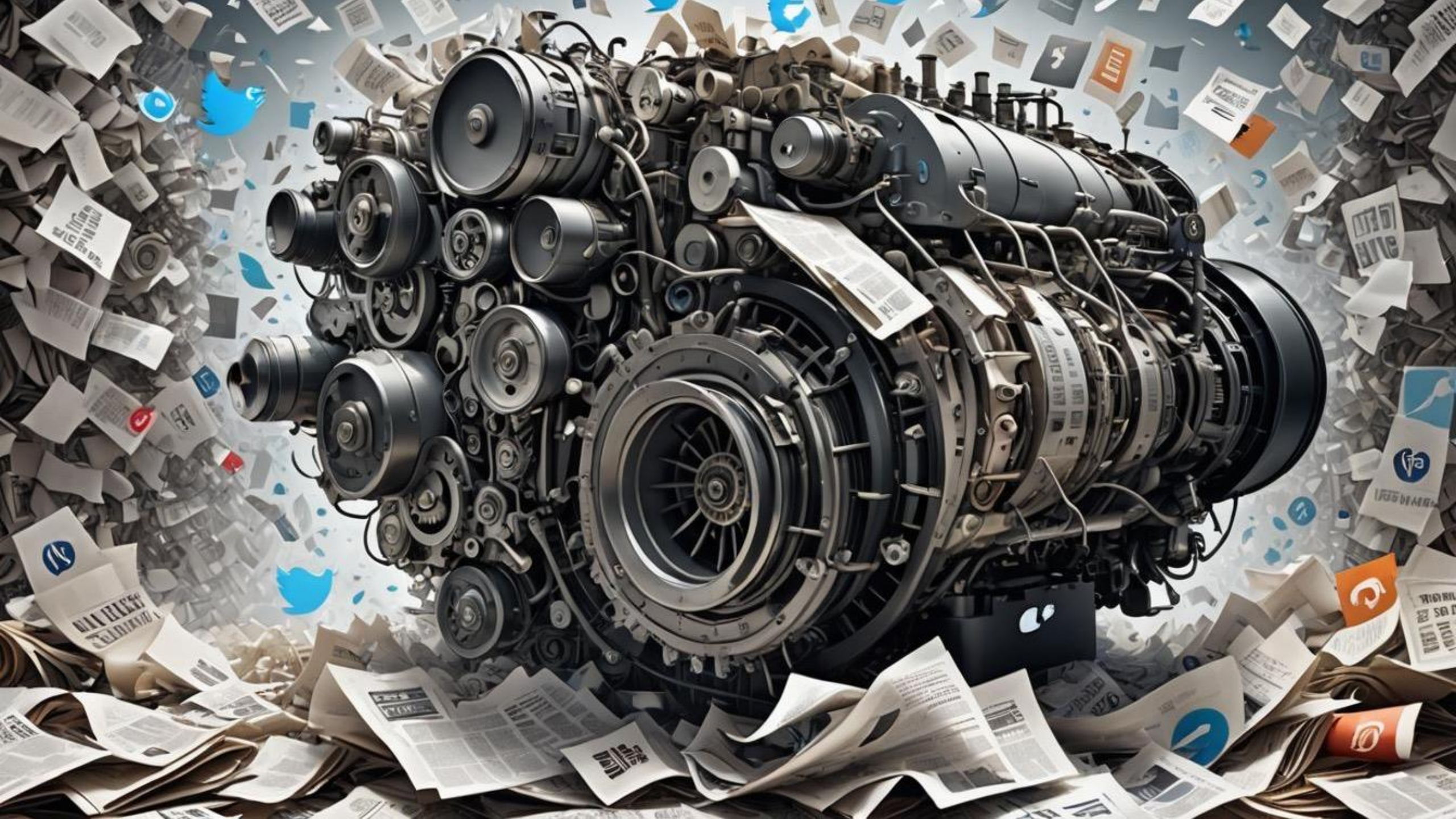


What kind of future did you think the internet would give us?

A Key Problem: Human Nature

- Early thinking about the internet underestimated critical facets of human nature, that have now come roaring forward.
- Personal confirmation and validation feel better than objective truth.
- There are lots of people in the fun house mirror business.
- You can create deceptions in the internet at all levels of abstraction.





Opinion | The internet was supposed to make humanity smarter. It's failing.



By Catherine Rampell
Columnist | + Follow
March 29, 2024 at 4:46 p.m. EDT

Influencer pricing: how much influencers really cost

by Sam Kendall | May 30, 2024 | 10 minutes

The impact of influencer marketing | Types of influencer partnerships | How much does influencer market



A growing number of Americans are getting their news from TikTok



Illustration by Nick Barclay / The Verge

MOTHERBOARD
TECH BY VICE

AI Trained on 4Chan Becomes 'Hate Speech Machine'

After 24 hours, the nine bots running on 4chan had posted 15,000 times.

A Russian Facebook page organized a protest in Texas. A different Russian page launched the counterprotest.

POLITICS OCTOBER 19, 2023

Media Confidence in U.S. Matches 2016 Record Low

BY MEGAN BRENAN



Deceptive Patterns

Home ›

Hall of shame

Hundreds of examples of deceptive patterns us



Enshittification

Enshittification is a pattern where online services and products experience a decline in quality over time.
Wikipedia

whoami



Tom Cross (Decius)

- Guy known for giving rants at hacker cons
- Creator of **FeedSeer**, a news reader for Mastodon, and **MemeStreams**, an early social network
- Security researcher (**IBM X-Force, Lancope**), CTO (**Drawbridge Networks, OPAQ, Fruitful**)
- Principal at **Kopidion**



Greg Conti

- Developed and Taught Information Operations course at BH and for enterprises
- **West Point, NSA, US Cyber Command**
- Dissertation was on Denial of Information Attacks
- Principal at **Kopidion**

We are teaching Adversarial Thinking at Defcon Training

The views in this talk are those of the speaker and do not reflect the policy or position of our current or former employers.

What this talk will cover

- **Deception:** A formal understanding of how to craft effective deceptions and deception operations, grounded in military doctrine.

- **Counter-Deception:** Disciplines that can be constructed within information or intelligence collection efforts that protect against deception.

- **Internet Applications:** How can we apply counter-deception disciplines to the design of capabilities within the internet that make deception harder.

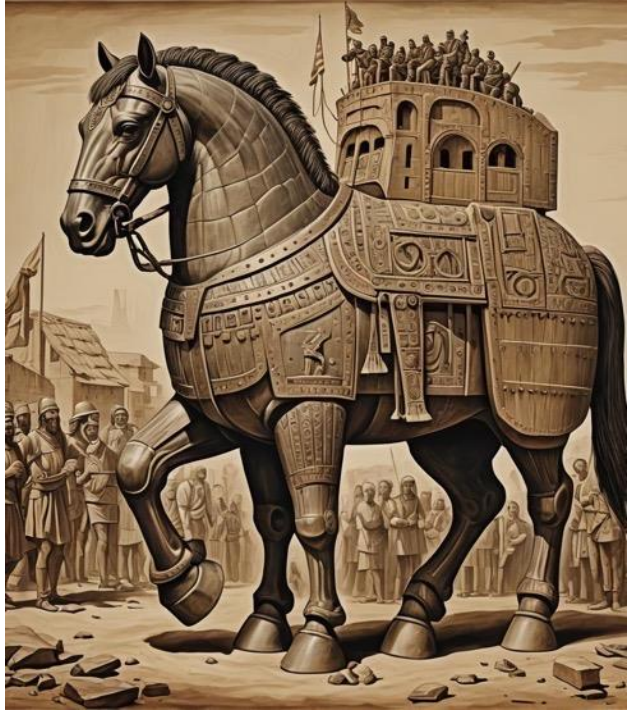


Why, So What, Who Cares

- Hackers have the ability to see things (fuckery) that other people don’t see
- If we don’t get this right – Now – we’ll be living in a future that out-Gibson’s William Gibson and out-Orwell’s George Orwell
- The people in this room have the skillset to deflect our future’s trajectory

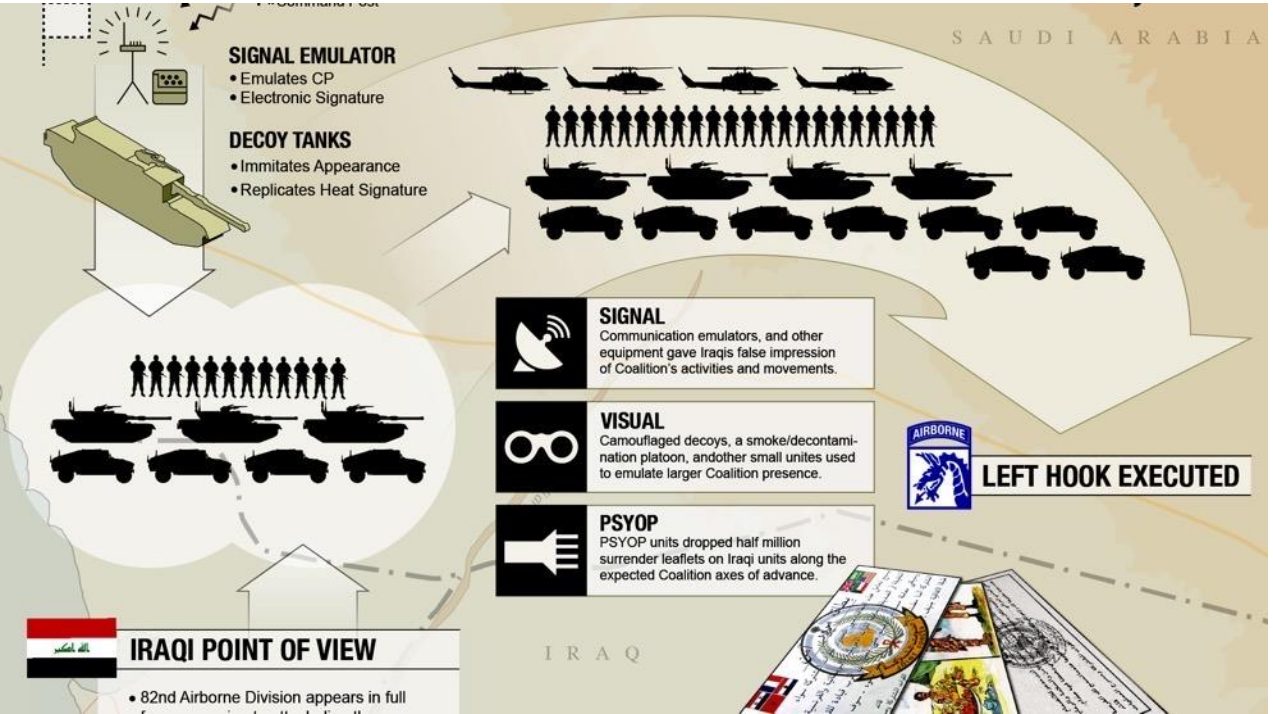












Deception is an Ancient Military and Intelligence Practice

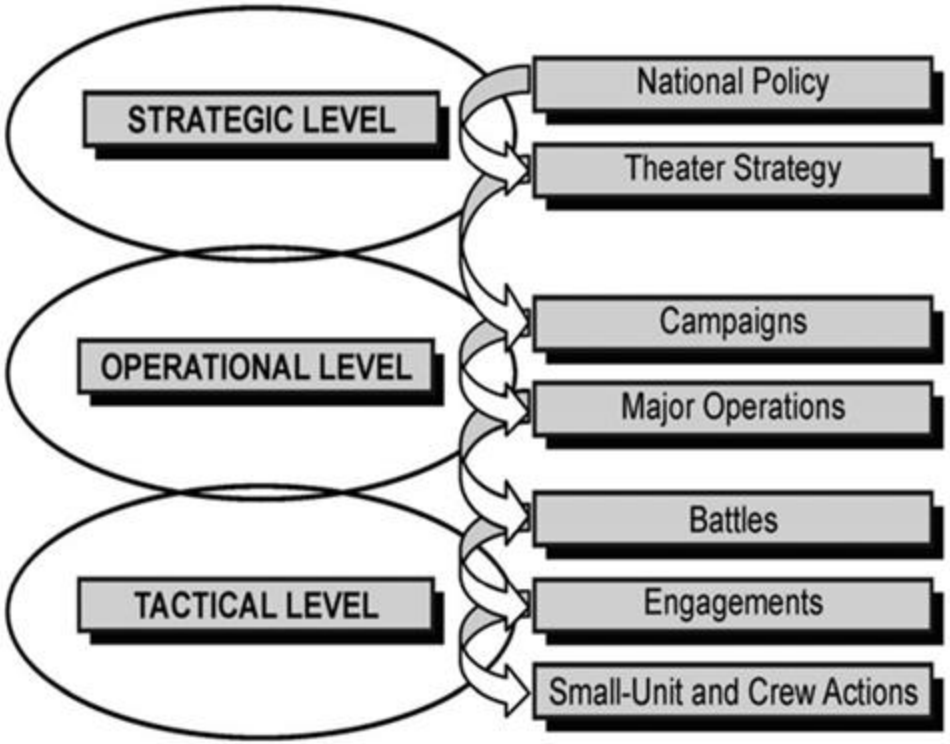
The act of hiding the truth, especially to get an advantage

Military deception is actions executed to deliberately mislead adversary military, paramilitary, or violent extremist organization decision makers, thereby causing the adversary to take specific actions (or inactions) that will contribute to the accomplishment of the friendly mission

https://dictionary.cambridge.org/us/dictionary/english/deception
https://www.cia.gov/resources/csi/static/Soviet-Deception-Cuban-Missile.pdf
https://en.wikipedia.org/wiki/Ghost_of_Kyiv#/media/File:Ghost_of_Kyiv_Art.jpg
https://www.armyupress.army.mil/Books/Browse-Books/iBooks-and-EPUBs/Deception-in-the-Desert/

	Deception Target			
	Human		Code	AI
	Users	Malware Analysts	Malware Detection Systems	AI
Deception Technique	Phishing Typosquatting Domain mimicry Spoofed login page Malicious PDFs Malicious Images Trojans Keyloggers	False Flags Code obfuscation/packing Fileless malware Deceptive metadata Anti-debugging Living off the Land Attacks C2 on legitimate domains	Fileless Malware Polymorphic malware Stolen digital certificates Rotating C2 infrastructure Code injection Encryption Bypassing AV Impersonated system context	Poisoning Jailbreaks Evasion attacks Privacy attacks
Examples	My name is Legion; for we are many :) Zeus (banking)	Olympic Destroyer xz backdoor Petya & NotPetya CyberCaliphate & TV5Monde	Malvertising App Stores infiltration Antivirus software bypass	Glaze Nightshade

Levels of Deception



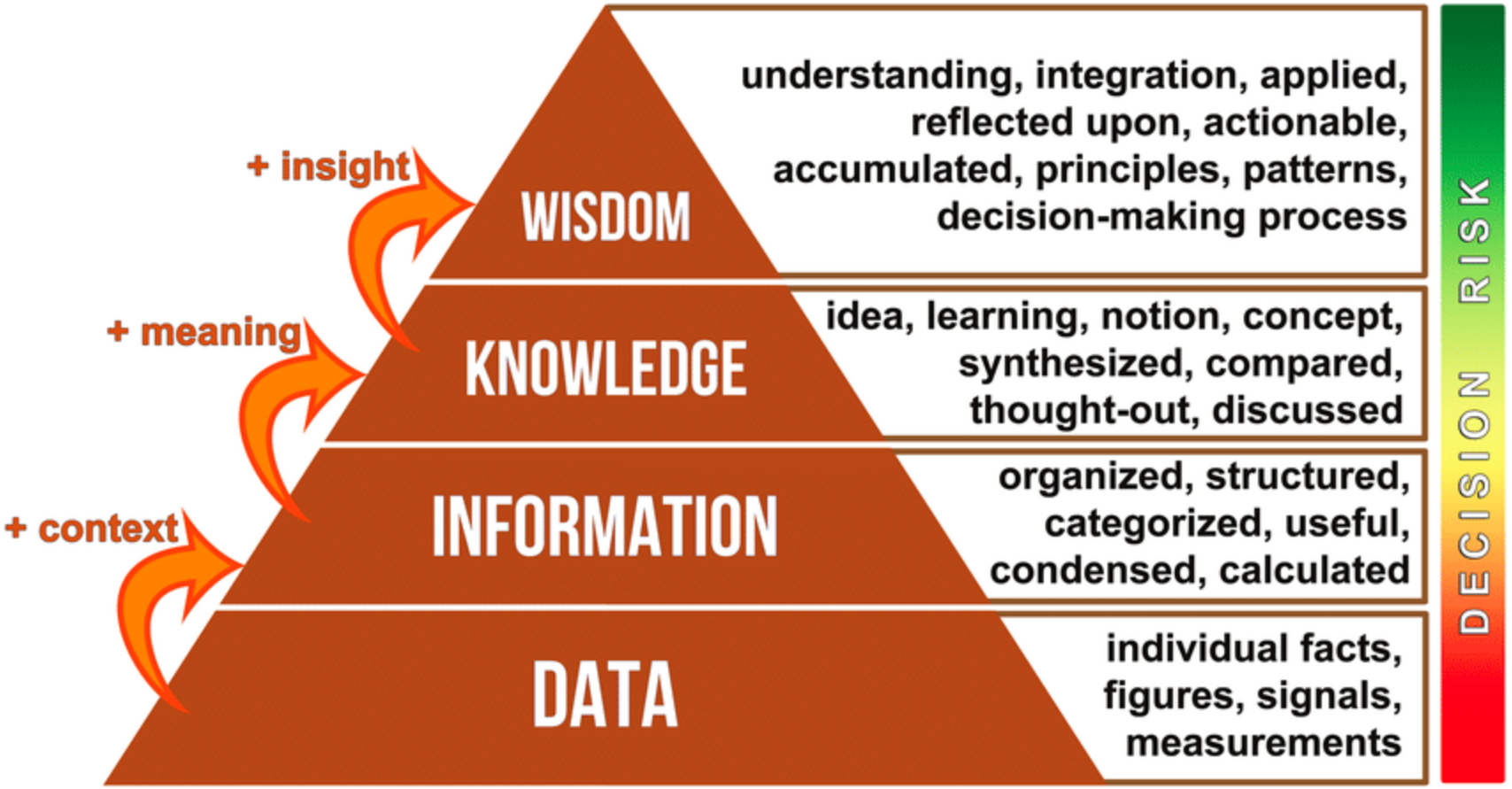
Strategic: Disguises basic objectives, intentions, strategies, and capabilities.

Operational: Confuses an adversary regarding a specific operation or action you are preparing to conduct.

Tactical: Mislead others while they are actively involved in competition with you, your interests, or your forces.

JW Caddel, Deception 101 - Primer on Deception, Strategic Studies Institute. At http://www.strategicstudiesinstitute.army.mil/pdffiles/pub589.pdf

Poisoning the DIKW Hierarchy



What are you really trying to do

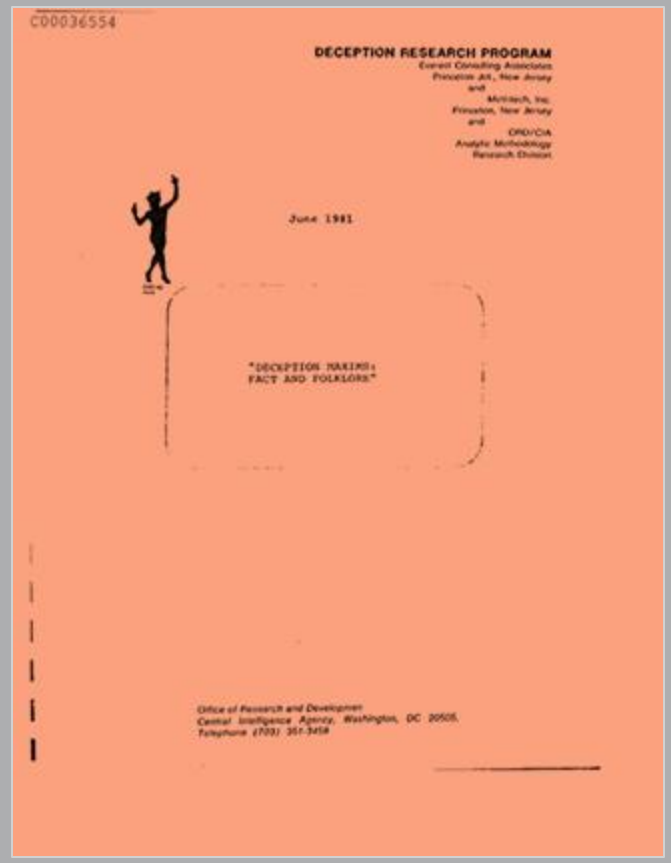
How deep are trying to get into your target's head

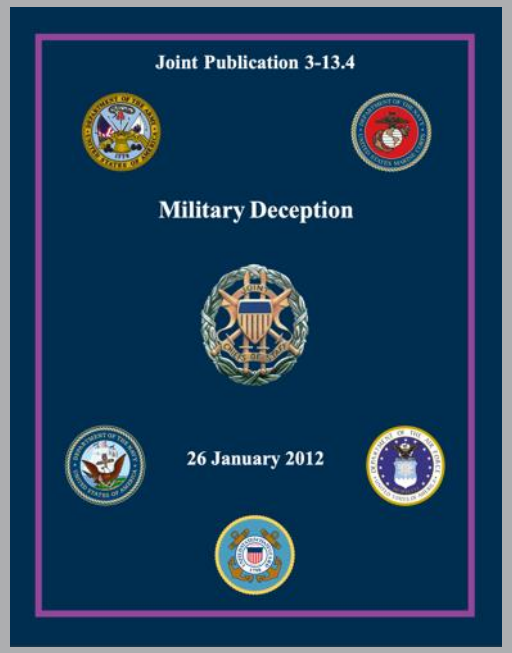
AI intersection?
The strategic high level perspective

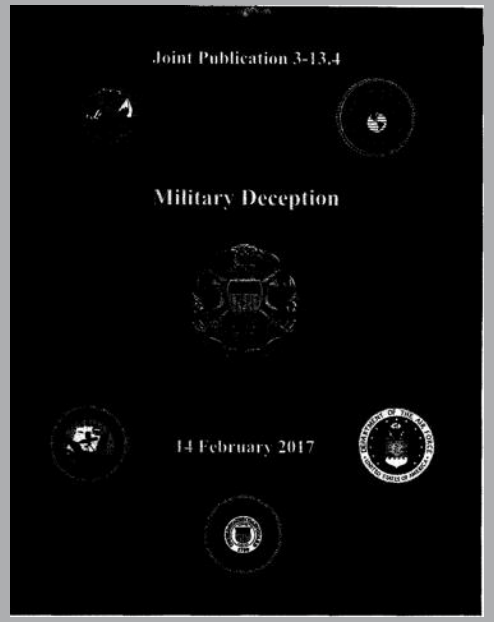
Image: Luis Tedeschi (2019), released under CC 4.0

Deception Maxims

- These deception maxims were developed by the CIA, later declassified and released under FOIA
- In this section of the talk, we explain these maxims and illustrate them with examples seen today

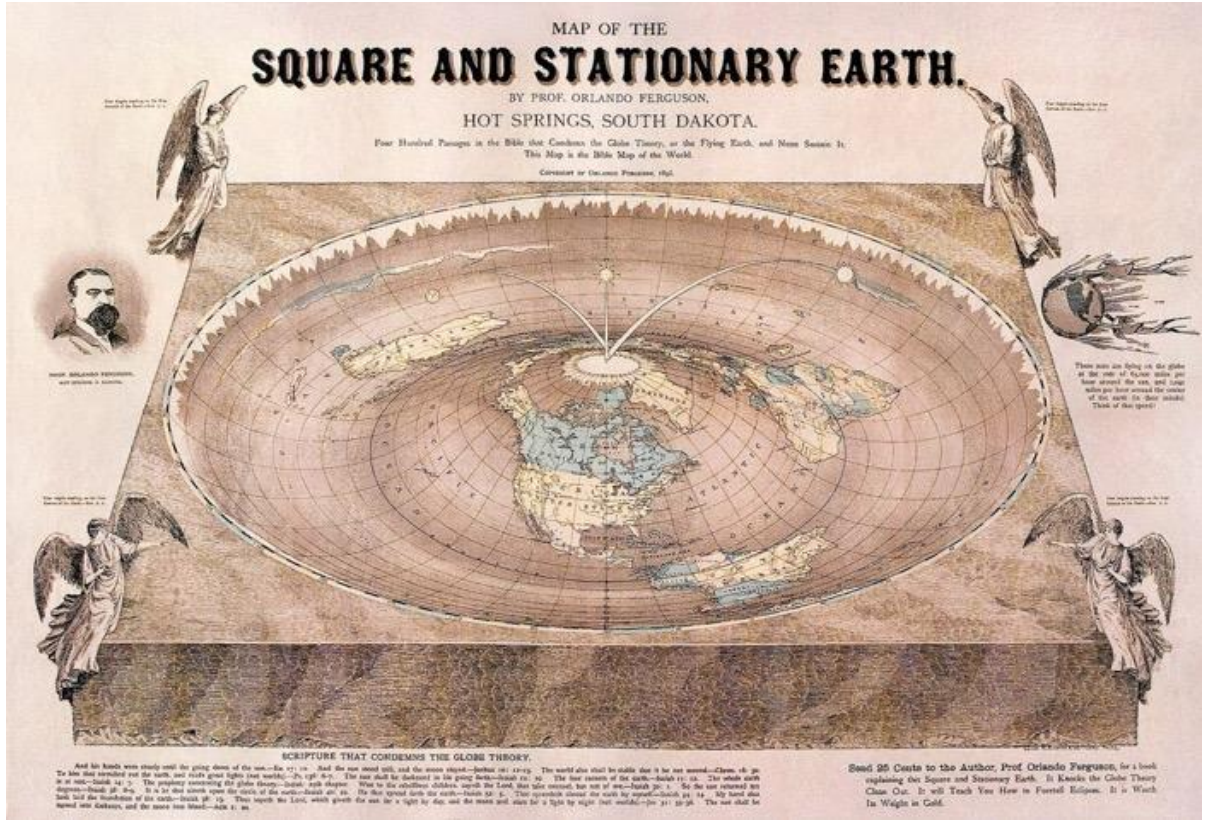






Maxim 1 - Magruder's Principle

It is easier to maintain a pre-existing belief in the target than force a change in beliefs





"I do have the ability to detect and affect human emotion...

When I do interfere, it is rarely that I do more than strengthen, as little as I can, what is already there...

He was not anxious to hand you over to the Imperial authorities, whom he does not like. I merely strengthened the dislike a trifle and he handed you over to my care, accepting the arguments I offered..."

Asimov – Prelude to Foundation

Maxim 2 - Exploiting Limits of Human and Machine Sensing and Information Processing

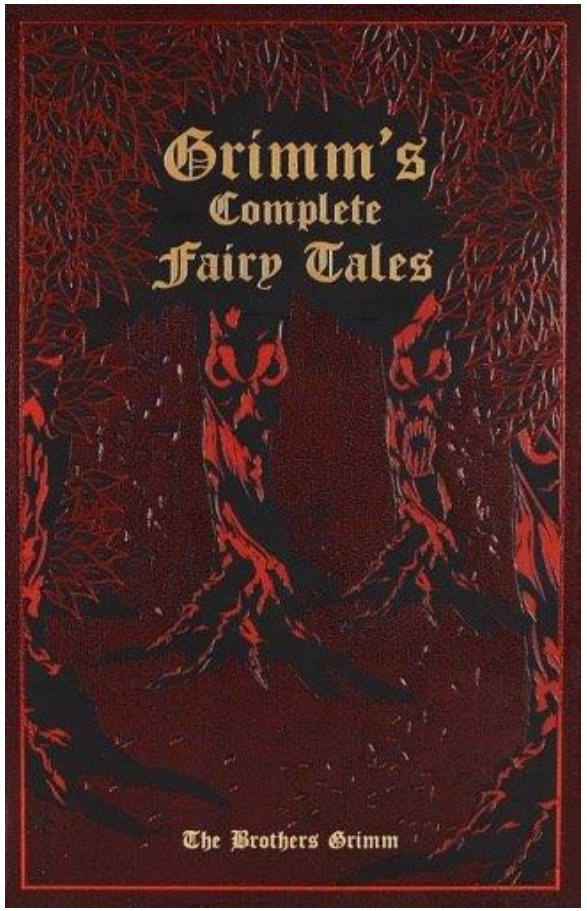


Maxim 3 - Jones Dilemma

Deception is difficult if there are more real sources than false sources



Maxim 4 - Carefully Sequence Deception Activities to Tell a Story



Maxim 5 - Carefully Design Planned Placement of Deceptive Material

- Make target “work” for it, don’t boldly announce what you are doing (example diary easily found on a desk vs. hidden, obfuscated code)
- Contrast this with “orgy of evidence”

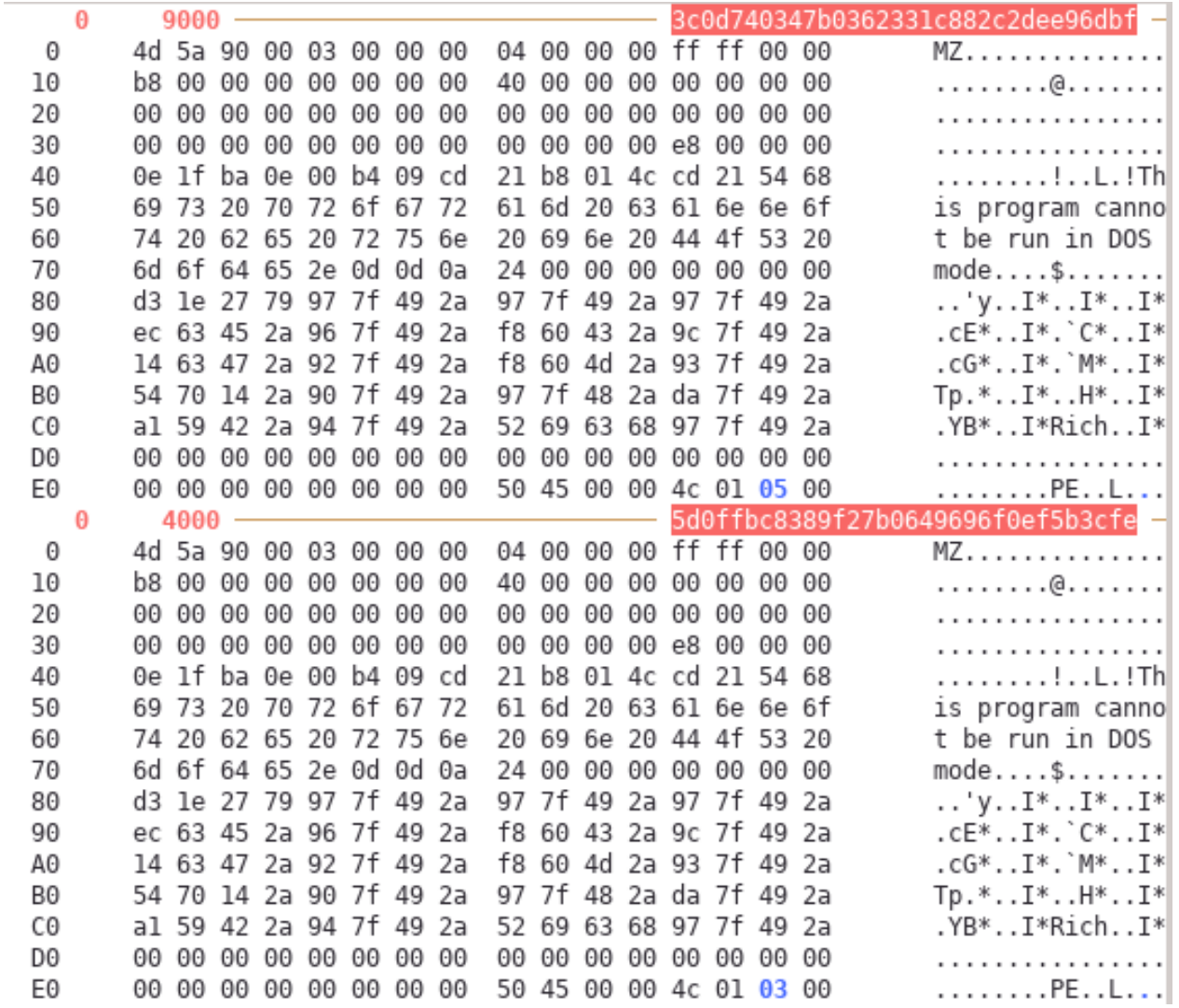


"A combination of certain code development environment features stored in executable files, known as Rich Header, may be used as a fingerprint identifying the malware authors and their projects in some cases.

In case of Olympic Destroyer wiper sample analyzed by Kaspersky Lab this “fingerprint” gave a 100% match with previously known Lazarus malware components and zero overlap with any other clean or malicious file known to date to Kaspersky Lab.

It now looks like a very sophisticated false flag which was placed inside the malware intentionally in order to give threat hunter impression that they found a smoking gun evidence."

- Kaspersky Labs



https://securelist.com/olympicdestroyer-is-here-to-trick-the-industry/84295/
https://securelist.com/the-devils-in-the-rich-header/84348/

Maxim 6a - Ambiguity

Increase doubt by providing many possible truths



Maxim 6b - Misdirection

Decrease doubt by focusing target on a given falsehood



https://commons.wikimedia.org/wiki/File:V%C4%9Btru%C5%A1e,_zrcadlov%C3%A9_bludi%C5%A1t%C4%9B.jpg

Maxim 7 - Husband Deception Assets



Maxim 8 - Feedback



Includes attackers monitoring for success of deception and monitoring for being deceived themselves.

Countering Deception

Four ways to counter deception:

- **Intelligence:** Directly observing the adversary's deception planning and activities
- **Disruption:** Preventing the adversary from injecting narratives into your sources
- **Analytic:** Detecting simulated narratives
- **Deterrent:** Demonstrating detection capability

Developing Analytic Processes:

- Identifying and checking key assumptions
- Devil's Advocacy or TeamA/TeamB
- What do deception maxims imply about defense?



Maxims that suggest when deception may be present

Maxim	Counter-Maxims
Maxim 6a - Ambiguity Deception	Is there more than one narrative present? Contradictions and incongruities may suggest deception.
Maxim 6b - Misdirection Deception	The Plus-Minus Rule: No imitation can be perfect without being the real thing. In a simulation, characteristics will be added or they will be missing.
Maxim 4 - Carefully Sequence Deception Activities to Tell a Story	Is there an "orgy of evidence?" Are facts being revealed all at once when it would make more sense for them to emerge slowly (or vice-versa).

Maxims that suggest mental discipline in analysis

Maxim	Counter-Maxims
Maxim 1 - Magruder's Principle	Apply the same critical analysis to facts that support your assumptions that you apply to facts that contradict them.
Maxim 5 - Carefully Design Planned Placement of Deceptive Material	Don't assume a fact is true just because it was hard to obtain it.

"The vulnerable mind fits ambiguous information to its own preconceptions and expectations."

- Bennet & Waltz - CounterDeception Principles and Applications for National Security



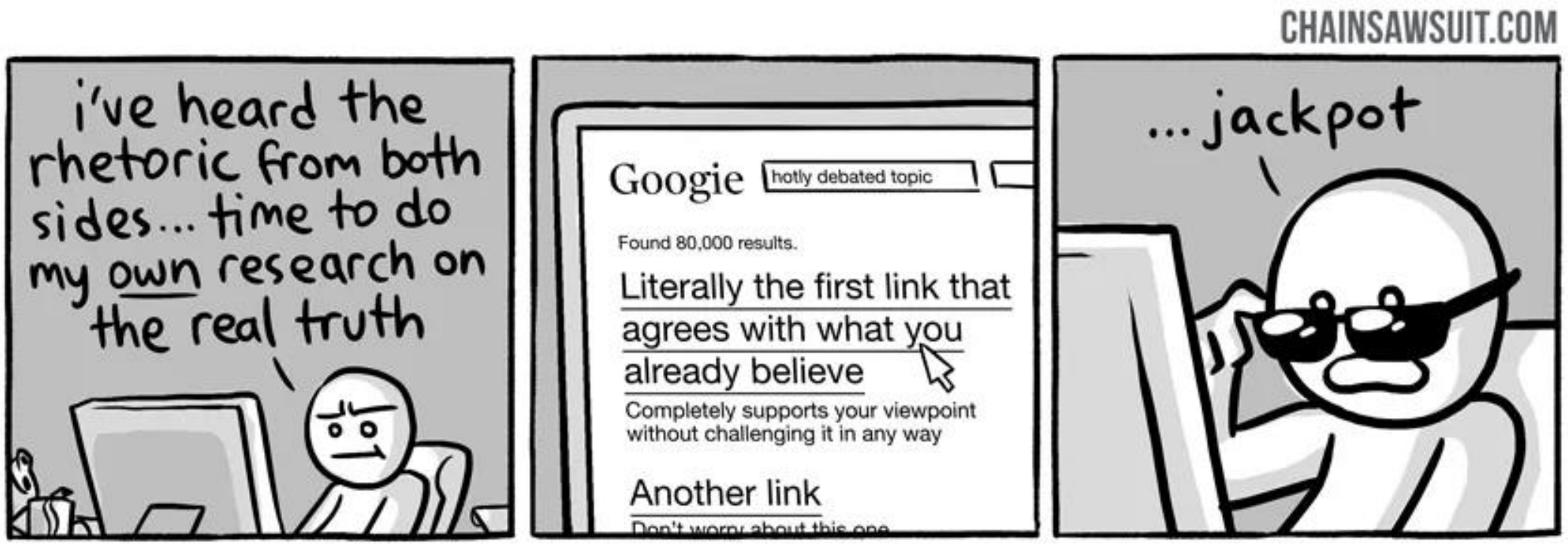
https://en.wikipedia.org/wiki/Media_bias
https://www.hunimed.eu/news/list-of-common-cognitive-biases/

Maxims that suggest methods of preventing deception

Maxim	Counter-Maxims
Maxim 2 - Exploiting Limits of Human and Machine Sensing and Information Processing Maxim 3 - Jones Dilemma	Develop lots of trustworthy sources of data and analysis. Measure reality from multiple perspectives and resolutions.
Maxim 7 - Husband Deception Assets	Model the adversary's capabilities and neutralize them.
Maxim 8 - Feedback	Behave as if you believe the deception so the adversary doesn't pivot. Deter the adversary by showing them that you've detected the deception. (Pre-bunking)

Developing trustworthy sources of data and analysis

- Evaluating independence, expertise, and the biases of sources of analysis.
- **Do you trust this source because of your own biases?**
- Evaluating past performance to determine future results.
- Measuring from multiple perspectives and resolutions.
- Triangulating sources of information



https://chainsawsuit.krisstraub.com/20140916.shtml

A problem with experts.

We can't be experts on everything, but we have to make decisions about things.

We rely on experts to tell us both what the facts are and what to think about them, and that puts us at their mercy.

Experts don't have to lie – they can omit and emphasize.

Reality: 1 + 2 + 3 + 4 + 5 = 15

One Perspective: 1 + 2 + 3 = 6

Another Perspective 3 + 5 = 8



An observation about Journalistic Objectivity.

One Perspective: 1 + 2 + 3 = 6

Another Perspective 3 + 5 = 8

- Presenting both of these perspectives doesn't get us closer to the answer.
- The agenda setting power of institutions allows them to influence what questions are discussed. (See: Steven Lukes)

What if journalistic objectivity was defined as filling in the missing, relevant pieces of information that other institutions are overlooking?

"Good reporting requires the exercise of the highest of scientific virtues."

- Walter Lippmann

Reality: 1 + 2 + 3 + **4** + 5

Call to Arms

- Hackers have an independent view outside of traditional institutions, and can therefore imagine what is needed to empower individuals.

- Some of you remember Gopher - a past when the internet had other tools and other possibilities than it does today.

- Technology criticisms sometimes prompt overcorrection. Hackers often see both the promise and the perils.



Four Focus Areas.

Counter-Maxims	Projects
Model the adversary's capabilities and neutralize them.	DISARM Framework Misinformation Village
Measure reality from multiple perspectives and resolutions.	Can we build tools for information triangulation?
Apply the same critical analysis to facts that support your assumptions that you apply to facts that contradict them.	How do we know when questions have been raised about what we're reading?
Develop lots of trustworthy sources of data and analysis.	How can we identify and collect networks of experts?

WikiScanner and WikiWatchdog

WikiScanner

28 languages

Article Talk Read Edit View history Tools

From Wikipedia, the free encyclopedia

WikiScanner (also known as **Wikipedia Scanner**) was a publicly searchable database that linked anonymous edits on Wikipedia to the organizations where those edits apparently originated. It did this by cross-referencing the edits with data on the owners of the associated block of IP addresses, though it did not investigate edits made under a username. It was created by Virgil Griffith and released on August 13, 2007.[1][2]

WikiScanner



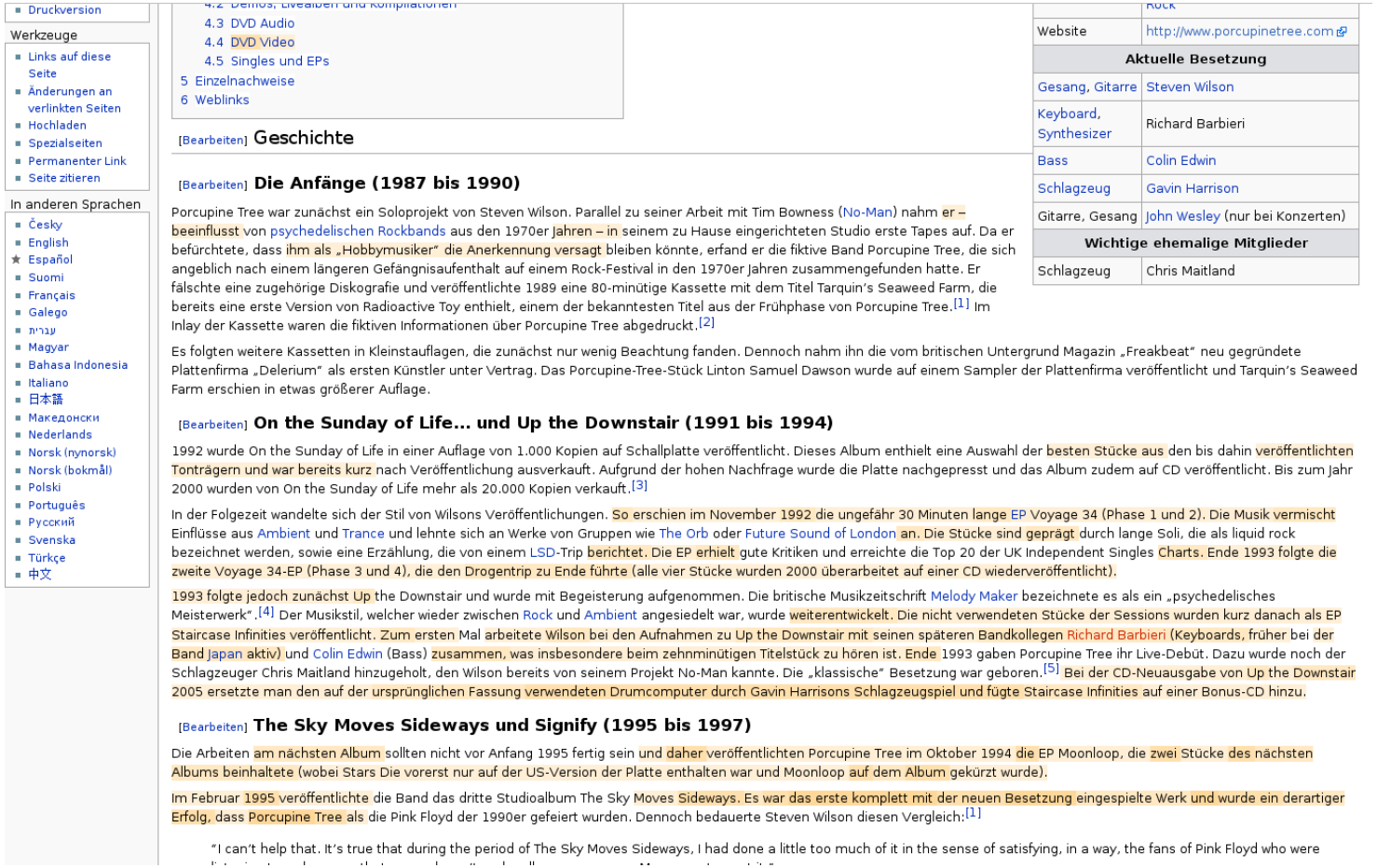
Screenshot of the website on August 22, 2007.

Type of site	Database tool
Available in	Chinese, Dutch, English, French, German, Italian,

Sometimes there are small, focused technical projects that can have a big impact.

https://github.com/volpino/WikiWatchdog

WikiTrust and Puppy Smoothies



- Druckversion

Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten
- Hochladen
- Spezialseiten
- Permanenter Link
- Seite zitieren

In anderen Sprachen
- Česky
- English
- ★ Español
- Suomi
- Français
- Galego
- עברית
- Magyar
- Bahasa Indonesia
- Italiano
- 日本語
- Македонски
- Nederlands
- Norsk (nynorsk)
- Norsk (bokmål)
- Polski
- Português
- Русский
- Svenska
- Türkçe
- 中文

4.2 Demos, Livealben und Kompilationen
4.3 DVD Audio
4.4 DVD Video
4.5 Singles und EPs
5 Einzelnachweise
6 Weblinks

	Rock
Website	http://www.porcupinetree.com
Aktuelle Besetzung	
Gesang, Gitarre	Steven Wilson
Keyboard, Synthesizer	Richard Barbieri
Bass	Colin Edwin
Schlagzeug	Gavin Harrison
Gitarre, Gesang	John Wesley (nur bei Konzerten)
Wichtige ehemalige Mitglieder	
Schlagzeug	Chris Maitland

[Bearbeiten] Geschichte

[Bearbeiten] **Die Anfänge (1987 bis 1990)**

Porcupine Tree war zunächst ein Soloprojekt von Steven Wilson. Parallel zu seiner Arbeit mit Tim Bowness (No-Man) nahm er – beeinflusst von psychedelischen Rockbands aus den 1970er Jahren – in seinem zu Hause eingerichteten Studio erste Tapes auf. Da er befürchtete, dass ihm als „Hobbymusiker" die Anerkennung versagt bleiben könnte, erfand er die fiktive Band Porcupine Tree, die sich angeblich nach einem längeren Gefängnisaufenthalt auf einem Rock-Festival in den 1970er Jahren zusammengefunden hatte. Er fälschte eine zugehörige Diskografie und veröffentlichte 1989 eine 80-minütige Kassette mit dem Titel Tarquin's Seaweed Farm, die bereits eine erste Version von Radioactive Toy enthielt, einem der bekanntesten Titel aus der Frühphase von Porcupine Tree.[1] Im Inlay der Kassette waren die fiktiven Informationen über Porcupine Tree abgedruckt.[2]

Es folgten weitere Kassetten in Kleinstauflagen, die zunächst nur wenig Beachtung fanden. Dennoch nahm ihn die vom britischen Untergrund Magazin „Freakbeat" neu gegründete Plattenfirma „Delerium" als ersten Künstler unter Vertrag. Das Porcupine-Tree-Stück Linton Samuel Dawson wurde auf einem Sampler der Plattenfirma veröffentlicht und Tarquin's Seaweed Farm erschien in etwas größerer Auflage.

[Bearbeiten] **On the Sunday of Life... und Up the Downstair (1991 bis 1994)**

1992 wurde On the Sunday of Life in einer Auflage von 1.000 Kopien auf Schallplatte veröffentlicht. Dieses Album enthielt eine Auswahl der besten Stücke aus den bis dahin veröffentlichten Tonträgern und war bereits kurz nach Veröffentlichung ausverkauft. Aufgrund der hohen Nachfrage wurde die Platte nachgepresst und das Album zudem auf CD veröffentlicht. Bis zum Jahr 2000 wurden von On the Sunday of Life mehr als 20.000 Kopien verkauft.[3]

In der Folgezeit wandelte sich der Stil von Wilsons Veröffentlichungen. So erschien im November 1992 die ungefähr 30 Minuten lange EP Voyage 34 (Phase 1 und 2). Die Musik vermischt Einflüsse aus Ambient und Trance und lehnte sich an Werke von Gruppen wie The Orb oder Future Sound of London an. Die Stücke sind geprägt durch lange Soli, die als liquid rock bezeichnet werden, sowie eine Erzählung, die von einem LSD-Trip berichtet. Die EP erhielt gute Kritiken und erreichte die Top 20 der UK Independent Singles Charts. Ende 1993 folgte die zweite Voyage 34-EP (Phase 3 und 4), die den Drogentrip zu Ende führte (alle vier Stücke wurden 2000 überarbeitet auf einer CD wiederveröffentlicht).

1993 folgte jedoch zunächst Up the Downstair und wurde mit Begeisterung aufgenommen. Die britische Musikzeitschrift Melody Maker bezeichnete es als ein „psychedelisches Meisterwerk".[4] Der Musikstil, welcher wieder zwischen Rock und Ambient angesiedelt war, wurde weiterentwickelt. Die nicht verwendeten Stücke der Sessions wurden kurz danach als EP Staircase Infinities veröffentlicht. Zum ersten Mal arbeitete Wilson bei den Aufnahmen zu Up the Downstair mit seinen späteren Bandkollegen Richard Barbieri (Keyboards, früher bei der Band Japan aktiv) und Colin Edwin (Bass) zusammen, was insbesondere beim zehnminütigen Titelstück zu hören ist. Ende 1993 gaben Porcupine Tree ihr Live-Debüt. Dazu wurde noch der Schlagzeuger Chris Maitland hinzugeholt, den Wilson bereits von seinem Projekt No-Man kannte. Die „klassische" Besetzung war geboren.[5] Bei der CD-Neuausgabe von Up the Downstair 2005 ersetzte man den auf der ursprünglichen Fassung verwendeten Drumcomputer durch Gavin Harrisons Schlagzeugspiel und fügte Staircase Infinities auf einer Bonus-CD hinzu.

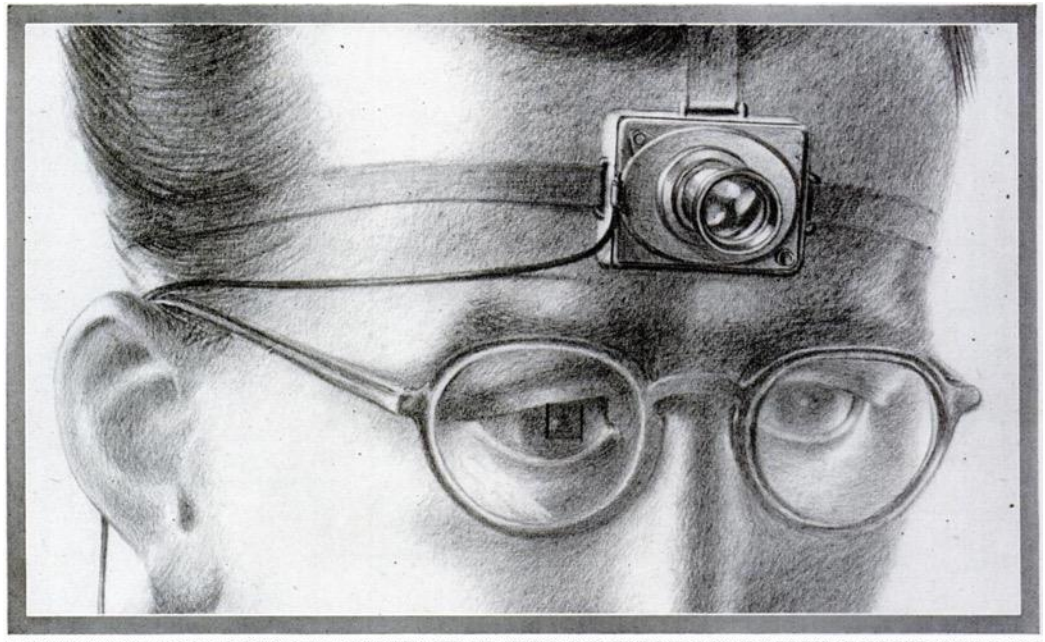
[Bearbeiten] **The Sky Moves Sideways und Signify (1995 bis 1997)**

Die Arbeiten am nächsten Album sollten nicht vor Anfang 1995 fertig sein und daher veröffentlichten Porcupine Tree im Oktober 1994 die EP Moonloop, die zwei Stücke des nächsten Albums beinhaltete (wobei Stars Die vorerst nur auf der US-Version der Platte enthalten war und Moonloop auf dem Album gekürzt wurde).

Im Februar 1995 veröffentlichte die Band das dritte Studioalbum The Sky Moves Sideways. Es war das erste komplett mit der neuen Besetzung eingespielte Werk und wurde ein derartiger Erfolg, dass Porcupine Tree als die Pink Floyd der 1990er gefeiert wurden. Dennoch bedauerte Steven Wilson diesen Vergleich:[1]

> "I can't help that. It's true that during the period of The Sky Moves Sideways, I had done a little too much of it in the sense of satisfying, in a way, the fans of Pink Floyd who were

Can we show the user when they are reading something that might be unreliable?

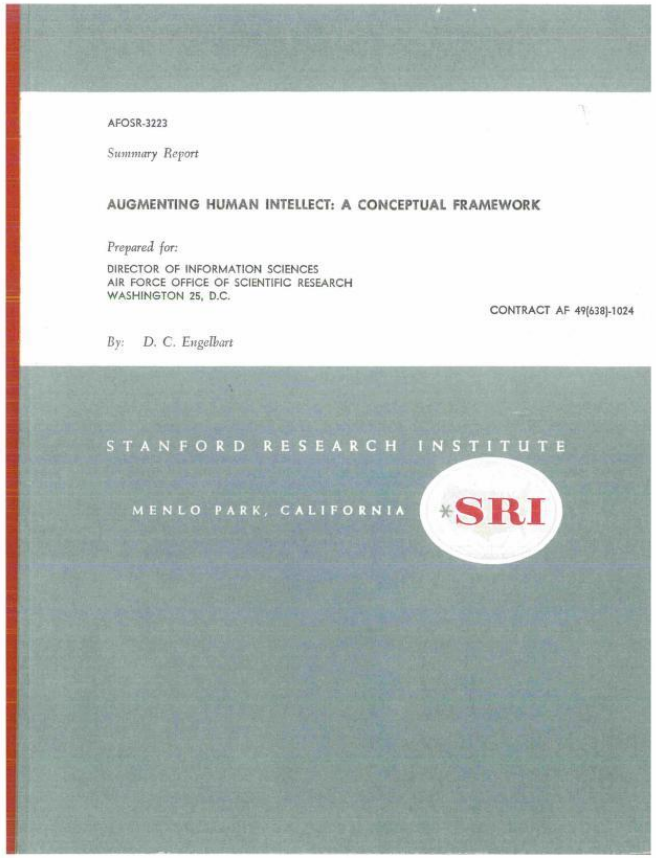


A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD IN WHICH MAN-MADE MACHINES WILL START TO THINK

Vannevar Bush



AFOSR-3223

Summary Report

AUGMENTING HUMAN INTELLECT: A CONCEPTUAL FRAMEWORK

Prepared for:
DIRECTOR OF INFORMATION SCIENCES
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
WASHINGTON 25, D.C.

CONTRACT AF 49(638)-1024

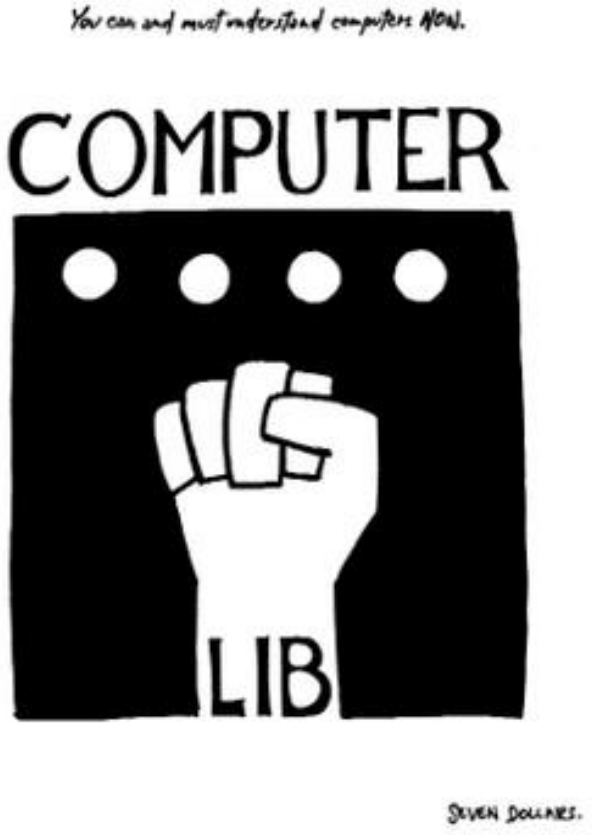
By: D. C. Engelbart

STANFORD RESEARCH INSTITUTE

MENLO PARK, CALIFORNIA

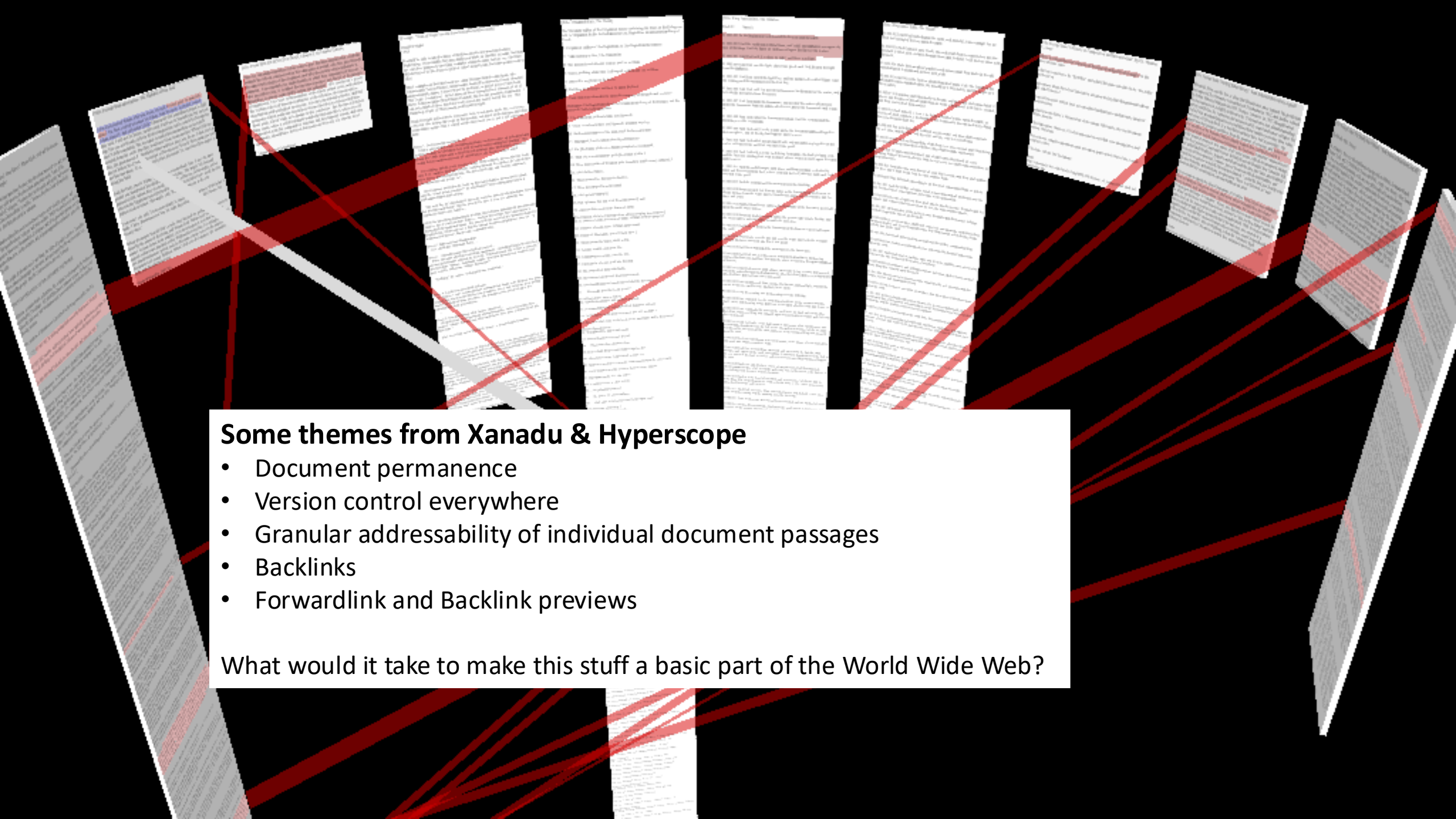
SRI

Douglas Engelbart



Ted Nelson

Theme: Computers will always augment human intelligence better than they can replace it.

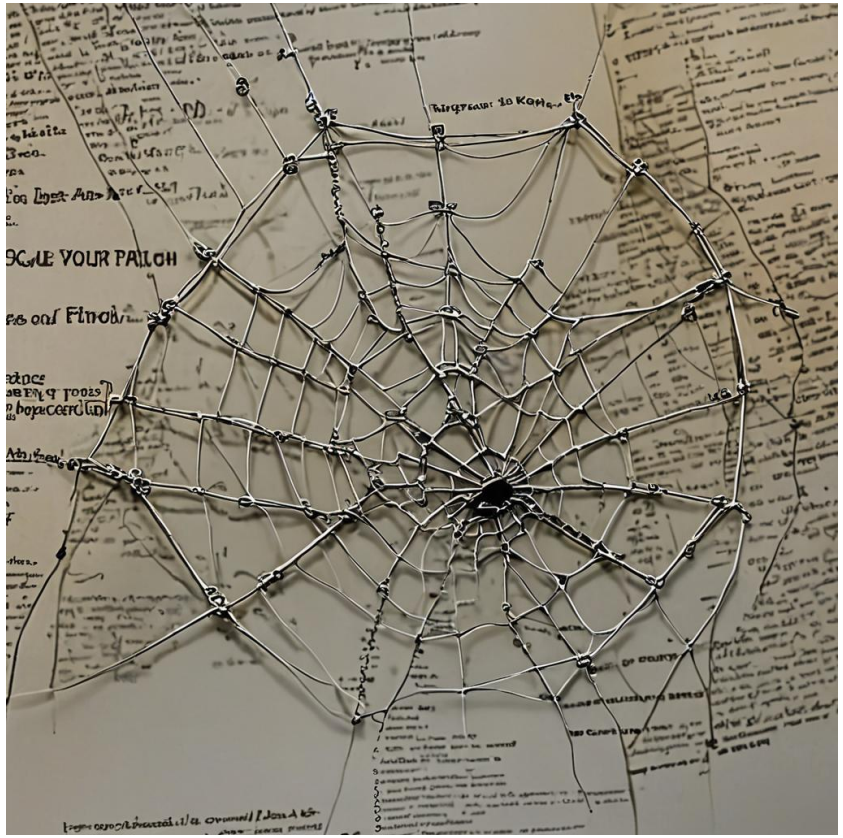


Some themes from Xanadu & Hyperscope

- Document permanence
- Version control everywhere
- Granular addressability of individual document passages
- Backlinks
- Forwardlink and Backlink previews

What would it take to make this stuff a basic part of the World Wide Web?

Backlinks are a content moderation challenge

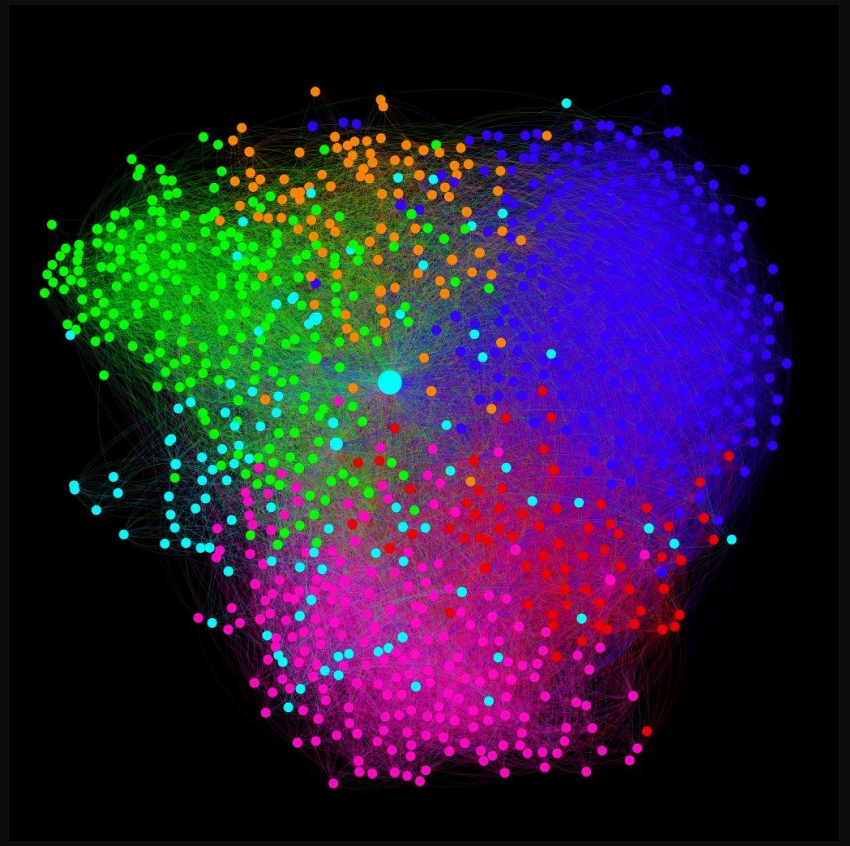


- Backlinks are both valuable and worthless
  - Pagerank uses backlinks to calculate a page's reputation
  - A naive open backlink system would be overwhelmed with abuse and spam
- Imagine if you had a single "comment section" that worked for every webpage on the Internet and only showed you posts from people whose views you respect and value.
  - Examples: rbutr, Genius Web Annotator, Google Sidewiki, hypothes.is
  - All opinions are not created equally - these systems must do more than provide annotations – they must surface the ones that are actually valuable.
- The people I have chosen to follow on my existing social media platform are people whose opinions are relevant to me, even if I don't often agree with them.
  - It would be valuable if my browser told me if anyone I currently follow on social media has posted about the page I'm viewing.
  - It might be even more valuable if social media users could respond to specific passages within web pages in a machine readable way.
- Building on top of social media platforms requires platforms that are technically, financially, AND culturally open to innovation.

Curating Collections of Experts: Expertise & Source Reliability Protocols

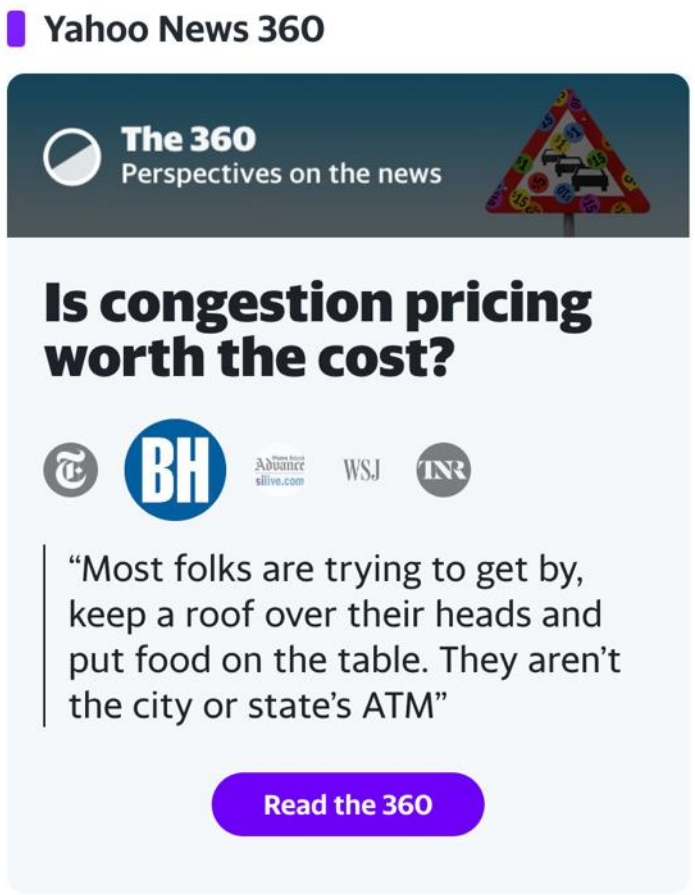
No current expertise infrastructure in the Internet combines:

- Open
- Machine readable
- Clear meaning & notion of authority



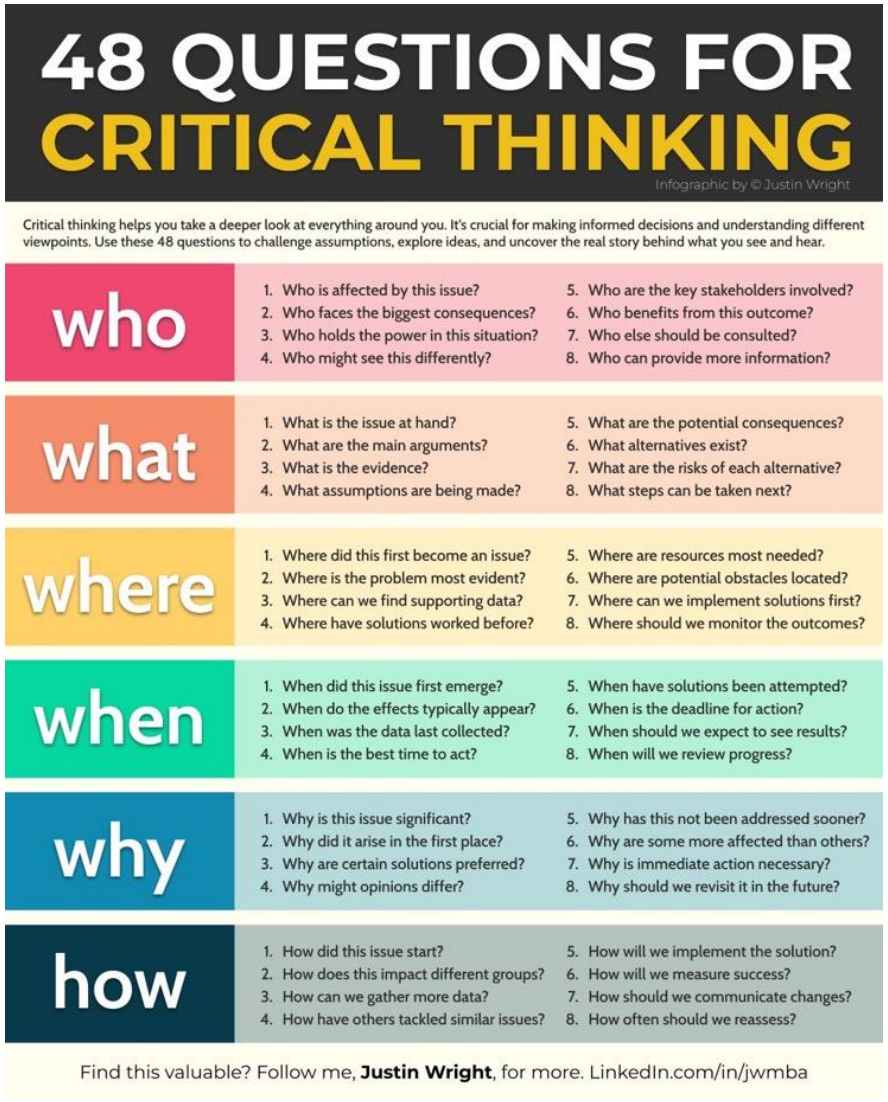
Authoritative: University Website Profiles
Quasi-Authoritative: Wikipedia Profiles
Web-of-Trust: Linkedin Endorsements
Self-Published: Social Media Profiles
Machine Readable: Geek Codes
Equivocal: Social Media Follower Graphs

Information Triangulation - Summarization of different views

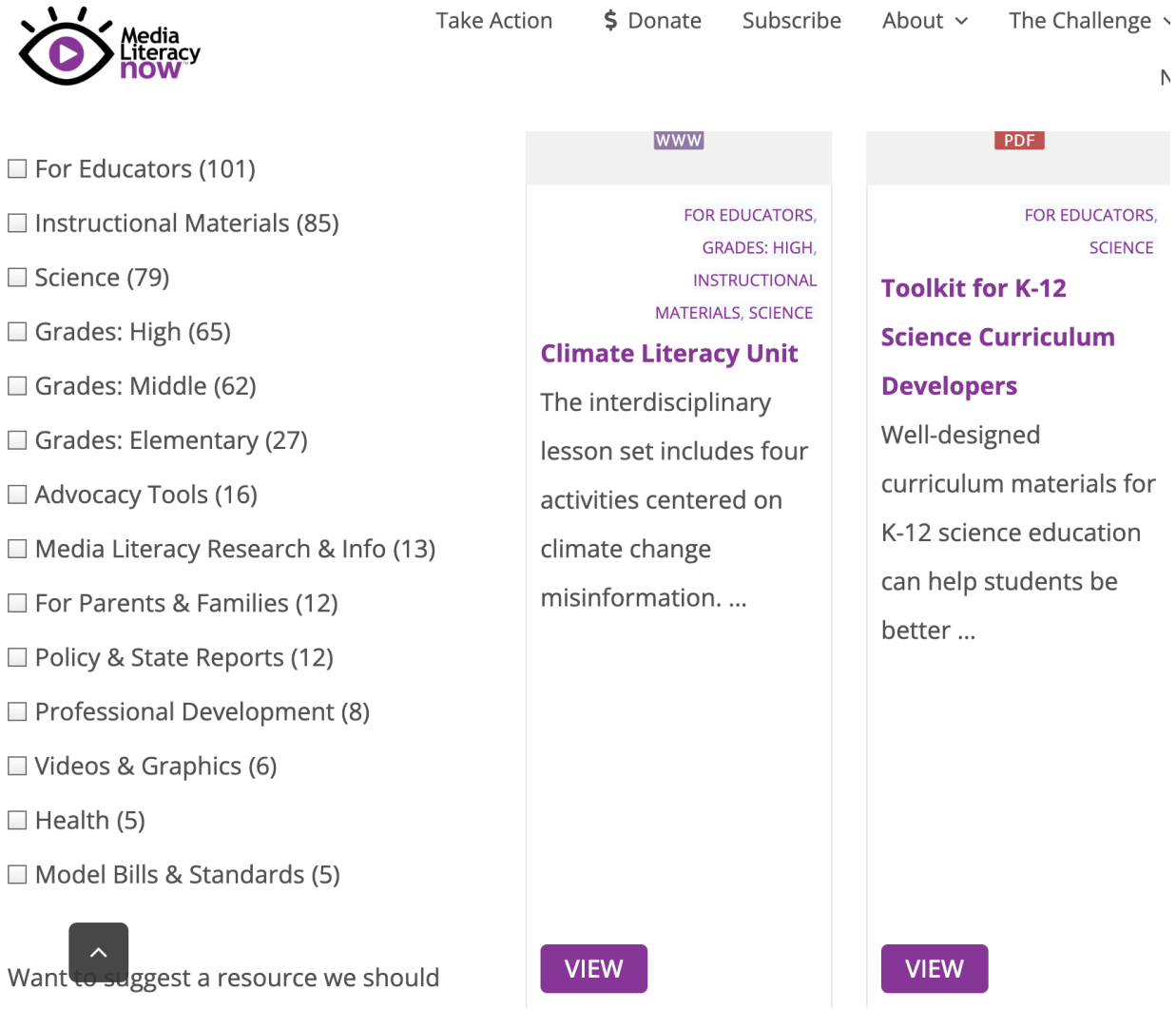


- Can I use LLMs to serve as a check and balance against my own biases?
  - LLMs are dispassionate
  - LLMs have built-in biases
- What if LLMs were coupled to data about source reputation?
- Could LLMs highlight individual passages in an article I am reading that have been disputed by other sources?
- Could LLMs scour university websites and produce machine readable indicators of academic's expertise?

https://news.yahoo.com/tagged/360/



Critical Thinking



Media Literacy

https://www.linkedin.com/posts/jwmba_90-of-people-lack-critical-thinking-skills-activity-7193961300748562432-exPW?utm_source=share&utm_medium=member_desktop

References

Deception:

- CIA Deception Maxims
- Army Support to Military Deception
- US DoD Doctrine on Military Deception

Deception in Malware:

- The Untold Story of the 2018 Olympics Cyberattack, the Most Deceptive Hack in History
- OlympicDestroyer is here to trick the industry
- Under False Flag: Using Technical Artifacts for Cyber Attack Attribution
- Wave Your False Flags! Deception Tactics Muddying Attribution in Targeted Attacks
- Digital False Flag Operations: A How-To Guide

Counter-Deception:

- Counterdeception Principles and Applications for National Security

Journalism, Media & Critical Thinking:

- 48 Questions for Critical Thinking - Justin Wright
- Media Literacy Now
- Journalism as a scientific endeavor - Julia Angwin
- How do we protect ourselves against disinformation? - Tom Cross
- Facts, frames, and (mis)interpretations: Understanding rumors as collective sensemaking – Kate Starbird

Internet Counter-Deception:

- DISARM Framework
- Misinformation Village
- RAND: Tools that Fight Disinformation Online
- The Socratic Web & Why Tools Shouldn't Adjudicate Truth - Shane Greenup

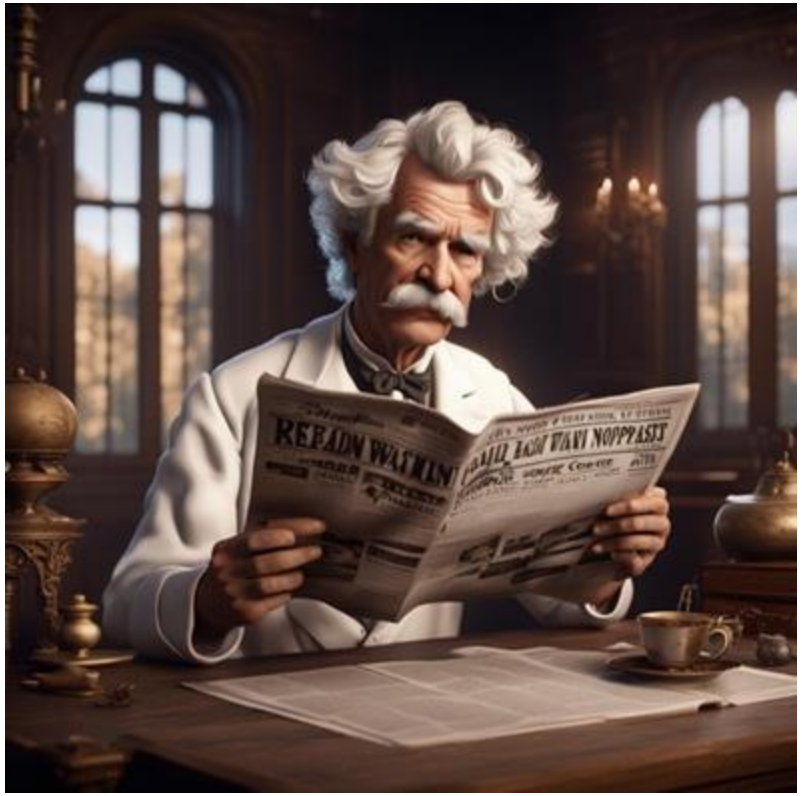
Wikipedia:

- WikiScanner & WikiWatchdog
- Puppy Smoothies (by Tom Cross) & WikiTrust

Inspiration for the World Wide Web:

- As We May Think - Vannevar Bush
- Augmenting Human Intellect – A Conceptual Framework - Douglas Engelbart
- Hyperscope 2.0
- Computer Lib/Dream Machines - Ted Nelson
- Project Xanadu

Thank you!



If you don't read the newspaper, you're uninformed. If you read the newspaper, you're mis-informed.

- Mark Twain

Tom Cross (aka Decius):

tom@tomcross.info

Greg Conti:

gjconti@rumint.org