Massimo Marchiori

Search Engines
and
Truth Manipulation

EU CULT

Brussels

22 Nov 2017

The Power of Search Engines

The main **gateway to information**

Primary role that naturally follows from
analysis of user behavior
($\rightarrow$ complementary to the social web)

Visibility

How important is the first page of results in a search engine?

First place: clicked by **51%** (!)

Top three: clicked by **73%** (!)

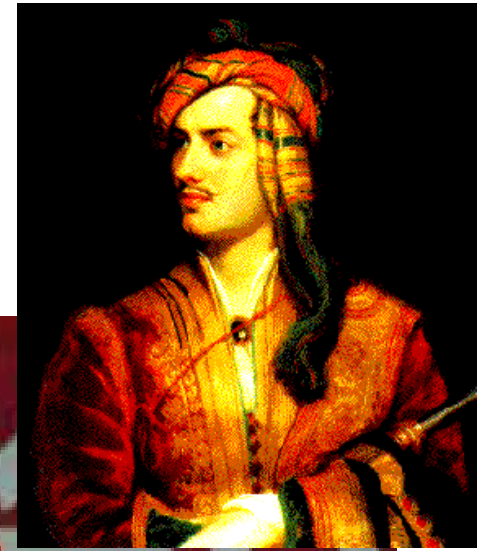Top ten: gets **95%** of all clicks (!)

→ *Visibility* is essential!

Resonation effects

Content is "resonated", and flows back and forth from search engines to the rest of web (including the social web)

Spyraling information that follows the **rich gets richer** phenomenon

Time stratification

"Fake news" isn't just it,
   "**fake information**" is broader

Time stratification = change things
   slowly, then let time incorporate those
   changes → bonuses by search engines

Using time stratification techniques, fake
   information can progressively change
   facts, history, and then trigger news

Example: Wikipedia
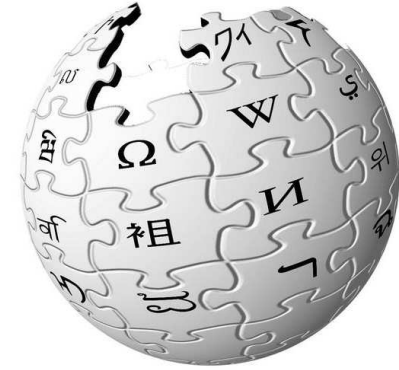
A primary target, both in itself, but also for the big amplification effect that creates via search engines

Wikipedia gets special priority by search engines, making it a primary target for *infiltration* of fake information of many kinds

Wikipedia → Negapedia

Negapedia is the ***negative version*** of Wikipedia

Exposes the huge "dark side", where people fight to change information, in battle rounds that evolve thru time

Thresholds and avalanches

The dangerous thing about propagation of fake news is that there **avalanches effects** (also called **cascades**)

After some **thresholds**, information spread in a booming way

The Nature of the Online World

The cascading phenomenon is due to structural properties of the web and of the social world

The online world is **super-connected** (small-world effect), and so fake information can propagate with incredible efficiency

The Crucial Zone

The crucial area for fake information is the ***initial phase*** of growth that can then trigger a cascade

Search engines are a primary way to trigger the avalanche (then expanding onto social sites like Facebook and Twitter)
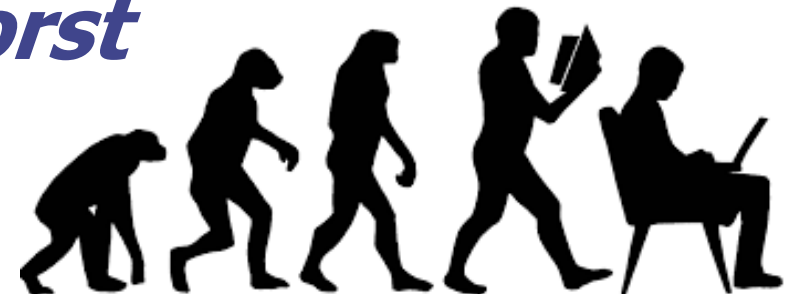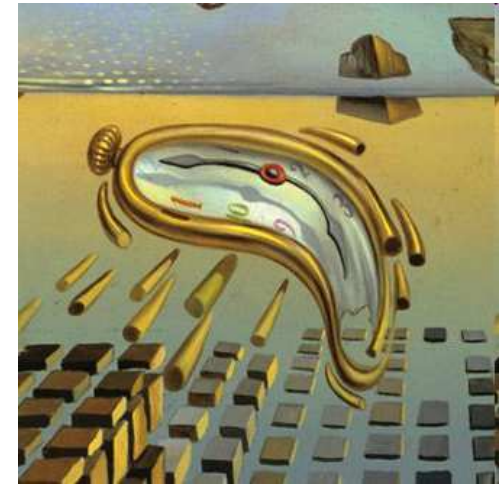
Where are we going?

Evolution Analysis

Takes into account various factors like the main *content amplifiers* (Google, Facebook, Twitter etc), the functional behavior of users, and the cost/benefit ratios

→ With the current settings, everything is getting *worst and worst*

Solutions?

Change the rules, change the cost/benefit
Two main complementary recipes for
search engines, valid also in general for
the main _content amplifiers_

First rule

Prioritize content safety

Currently, due to cost/benefit, content amplifiers try to *bypass responsability*

→ need instead to push content safety at top priority, not as a lateral annoyance

This is crucial to stop fake news already ***in the initial stage*** (otherwise cascades take place → ***too late*** !)

Second rule

Transparency

Now search engines and content amplifiers rule the kingdom of information

But in a peculiar way for a democracy: they mostly do it as they want (!)

The Black Box

They act as black boxes: rules governing information are secret

Dangerous!? Are they doing things impartially? Are they doing their best?

Shield: **Security by obscurity**

Rationale: need to "black box" the algorithms to protect us and you.

Incorrect!

Solution: Transparency

Proper transparency allows for proper **scrutiny** of the algorithms and policies in use within control amplifiers

This is critical: like in food processing, we need to be able to detect whether we are being served legal or poisoned food, as **information is the food of the third millennium** !