Science, Technology and Humanity: Opportunities and Risks

Future of minds, Transhumanism

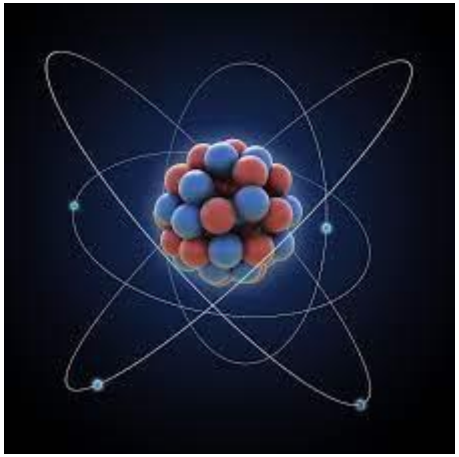
Martin Takáč, Tomáš Gál

http://dai.fmph.uniba.sk/courses/STH/

Nanotechnology

Nanotechnology

- The study, design, and manipulation of natural phenomena, artificial phenomena, and technological phenomena at the nanometer level. (Berne, 2005 in Tavani, 2013)
- 1nm = 10^{-9}m
- Close to the size of atom (0.1-0.5nm)



Existing nanotechnology research

- In 1990 Donald Eigler and Erhard Schweizer manipulated 35 individual xenon atoms to shape the IBM logo
- In 2007 nanotube radio invented by Alex Zettl et al.
- Hewlett Packard researchers have made computer memory devices by creating eight platinum wires 40 nm wide on a silicon wafer (1000 chips are the width of a human hair).

Predictions of future use

Optimistic view:

“Molecule by molecule manufacturing” would

- be self-sufficient and “dirt-free,”
- create unprecedented objects and materials,
- enable the production of inexpensive high-quality products,
- be used to fabricate food rather than having to grow it,
- provide low-priced and superior equipment for healthcare,
- enable us to enhance our human capabilities and properties.

(Gordijn 2003 in Tavani, 2013)

Predictions of future use

Pesimistic view:

- severe economic disruption
- premeditated misuse in warfare and terrorism
- surveillance with nano-level tracking devices
- extensive environmental damage
- uncontrolled self-replication (grey goo)
- misuse by criminals and terrorists (black goo)

(Gordijn 2003 in Tavani, 2013)

Ethical issues in nanotechnology

- privacy and control
- (negative effects of) longevity
- runaway nanobots

Question

Should nanotechnology research continue?

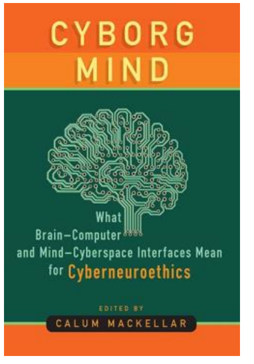
- View 1: if it is dangerous, the development should be restricted or stopped
- View 2: if research in nanocomputing and nanotechnology is prohibited/restricted, it will be done underground. Then nanotechnology research would not be regulated by governments and professional agencies concerned with social responsibility.
- View 3: The fact that it could cause damage doesn’t automatically mean it should not be developed (e.g. cars, medical drugs).

(Near) future of minds

Neuronal Interface Systems

1. **Interfacing out (output) of the nervous system**: this enables biological information to exit a neuronal network, such as the brain, which can then be sent to some form of computer that interprets the signal and triggers events or actions. For example, it enables brain information to be read and used in controlling a limb.
2. **Interfacing into (input) a nervous system**: this inputs information into a living neuronal network from outside, such as from a computer. For example, it enables a cochlear implant to provide sound information into the brain.
3. **Interfaces made of feedback loop systems**: these interpret information from a living neuronal network and sends it to an external processor, which then returns information back into the neuronal network.

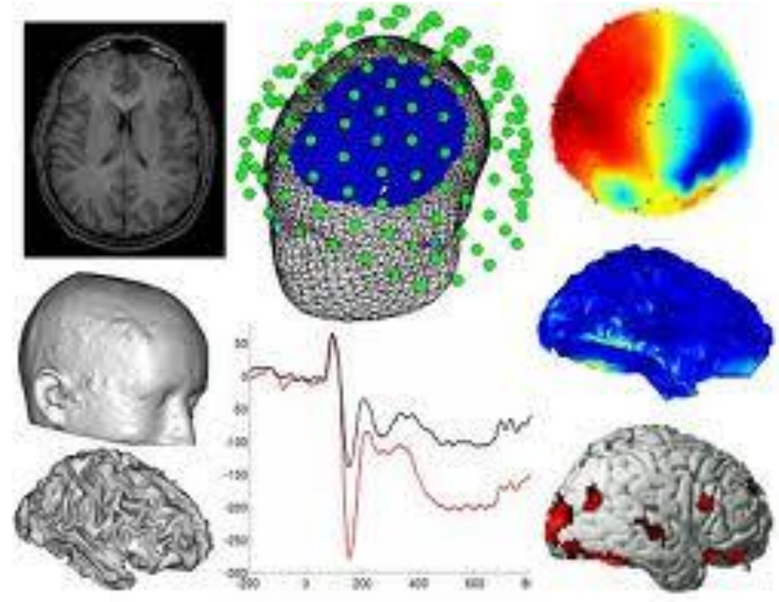
(MacKellar, 2019)



Output Neuronal Interface Systems

- **Invasive** – surgically implanted electrodes about 1.5–3 mm deep in the brain.
- **Partially invasive** – surgically implanted electrodes inside the scull, but on the brain surface
- **Noninvasive** – neuroimaging
  - Structural (anatomical) – brain architecture
  - Functional – brain activity

Neuroimaging



- X-ray
- Computed Tomography (CT)
- Positron Emission Tomography (PET)
- Magnetic Resonance Imaging (MRI)
- functional Magnetic Resonance Imaging (fMRI)
- Electroencephalography (EEG)
- Magnetoencephalography (MEG)
- Near-infrared spectroscopy (NIRS)
- Light and Electron Microscopy (LM and EM)

Invasive Input Neuronal Interface Systems

- neuronal implants for **deafness**
  - cochlear implants that bypass the dysfunctional signal recognition system in the ear
  - auditory brain stem implants that completely sidestep the whole hearing system.
  - Future:
    - integrate the implant with a mobile phone chip
    - or radio, TV, echolocation…
- Vision
  - Retinal implants for acquired blindness
  - Future:
    - wireless interfaces for TV or PC (viewing pictures)
    - Near-infrared vision for night driving

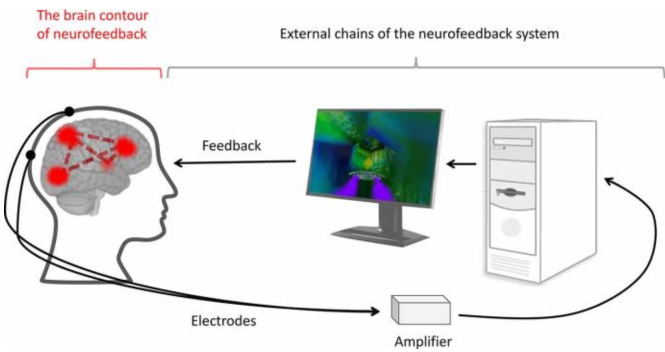
Invasive Input Neuronal Interface Systems

- Deep Brain Stimulation (DBS)
  - Long needles with electrodes on their tip sending pulses of electric current.
  - DBS shown to initiate metabolic and neurochemical brain changes when continual stimulation takes place
  - Applications:
    - Parkinson’s disease (increasing ability to move and walk).
    - Control of chronic pain, epilepsy, migraine, depression, Alzheimer’s disease and obesity
  - Future:
    - examine the effects of DBS on agency and decision-making because the procedure can directly change a person’s mood and behaviour by modifying the biological neuronal basis of unconscious and conscious mental states (e.g. for treatment of psychiatric disorders).

Noninvasive Input Neuronal Interface Systems

- Transcranial Brain Stimulation (TBS)
  - stimulate the brain either by inducing an electrical field using a magnetic coil placed against the head with transcranial magnetic stimulation (TMS) or by applying weak electrical currents via electrodes on the scalp with transcranial direct or alternating current stimulation (TDCS or TACS).
  - TMS approved in US for migraine and depression, considered for number of psychiatric and neurological disorders such as mania, obsessive-compulsive disorders, schizophrenia and Parkinson’s disease. TDCS for rehabilitation, depression and Parkinson’s disease treatment.
  - improved skill learning, multitasking and mood reported.

Feedback Systems of the Brain and Mind



Steps:

1. externalising brain activity (output);
2. pre-processing and making sure that background noises are addressed;
3. feature extraction that correlates brain signals to a small number of variables defined as features;
4. classification of the signals corresponding to a type of brain activity pattern;
5. translation into a command;
6. feedback in which a user is then informed of the brain activity that has been recognised.

Feedback Systems of the Brain and Mind

Applications:

- Neurorehabilitation
- Assistive technologies for people with disabilities (paralysis, spinal cord injury, stroke or amputation)
- Nerve recording implants (K. Warwick)
- Brain-computer interfaces (BCI) in gaming industry
- Brain decoding (“mind reading”)
- Future:
  - Brain controlled computers
  - Telepathy (brain-to-brain talk for military – DARPA)
  - Collective mind
  - Brain organoids grown synthetically in laboratories

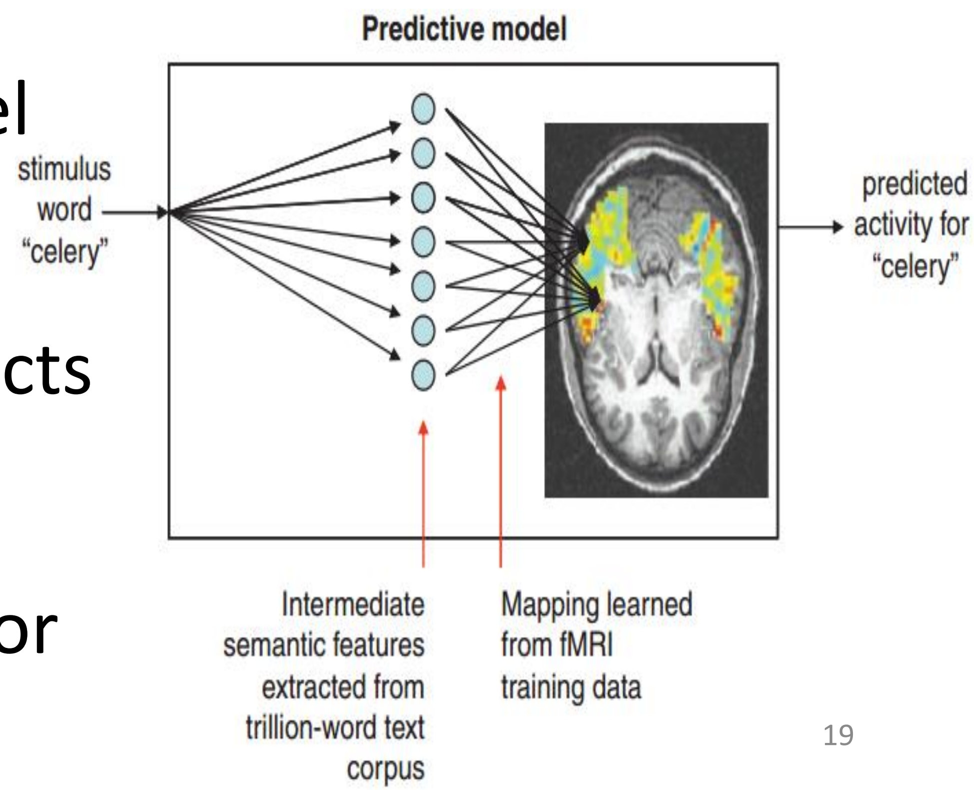
Can we use fMRI for mind reading?

- Mitchell et al, Science, 2008.
- Pereira et al, Frontiers in Human Neuroscience, 2011.
- Nishimoto et al, Cur. Biology, 2011.

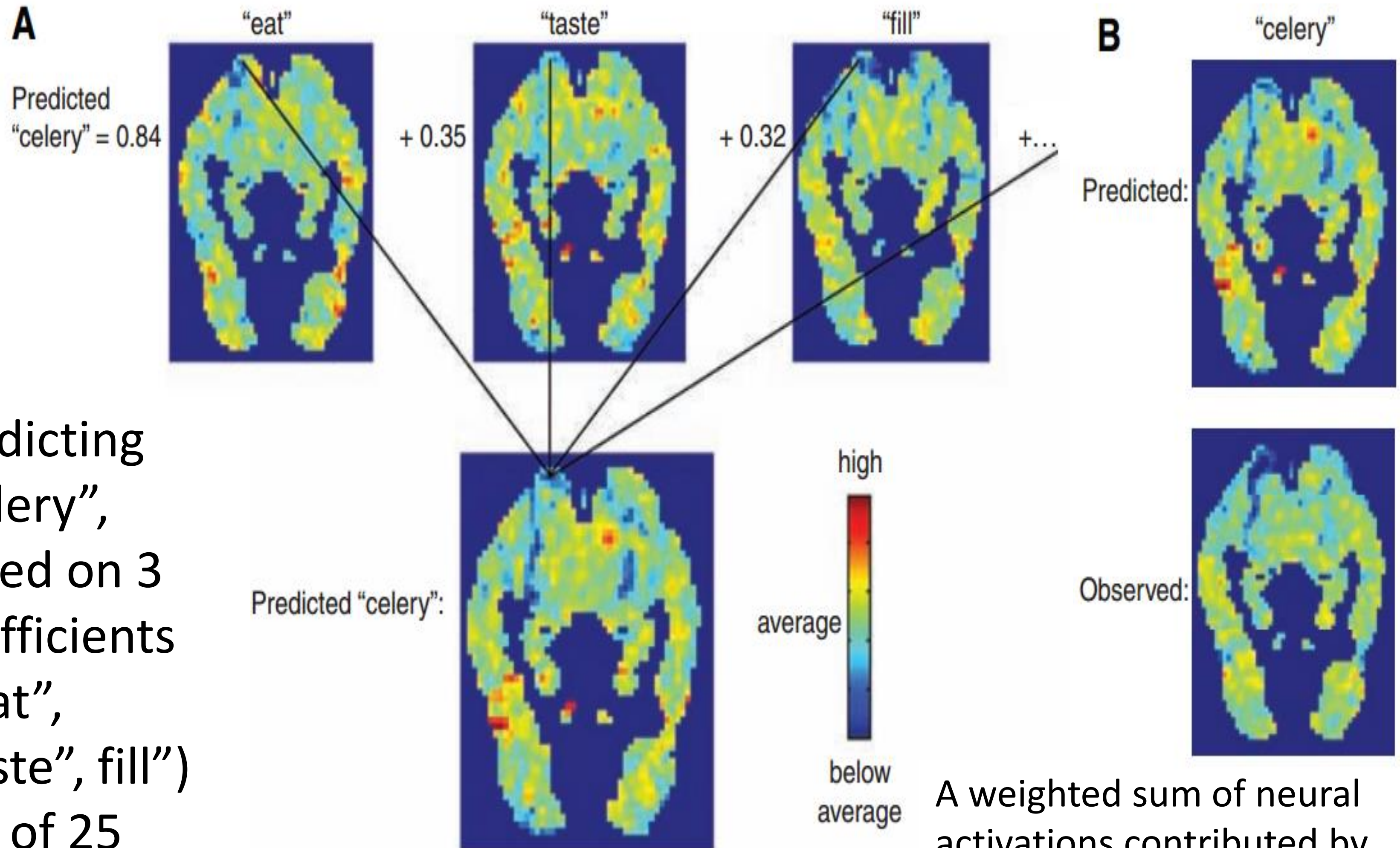
Predicting fMRI activation patterns

- Representation of concrete nouns is related to the distributional properties of those words in broadly based groups of the language

➔ A computational model trained to predict brain activity (fMRI) for the meaning of concrete objects (e.g. “celery”), even for objects that no fMRI activation patterns exist for



Example Prediction



Predicting "celery", based on 3 coefficients ("eat", "taste", fill") out of 25 semantic features

A weighted sum of neural activations contributed by each of these features is used to compute the fMRI activation at every voxel

Assumptions/Basic functionality

- meaning of the word is encoded as a vector of semantic features, based on the occurrences of that word within a (very large) corpus

- A weighted sum of neural activations contributed by each of these features is used to compute the fMRI activation at every voxel

- Model instantiated via a set of semantic features, “grounded” in sensory-motor actions (e.g. “hear”, “lift”, “fill”, “open”, etc.)

Model Performance

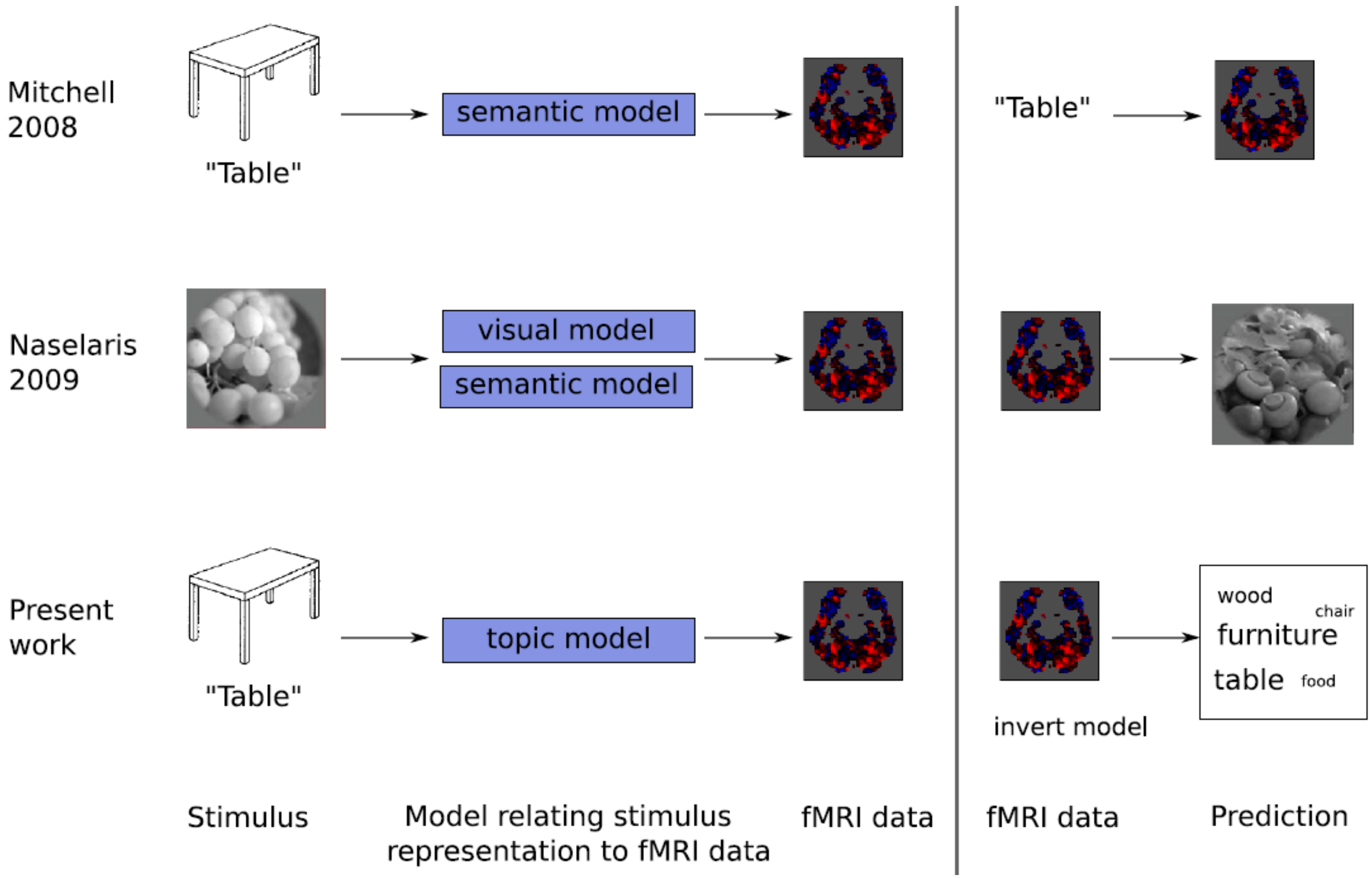
- Mean accuracy (training sets vary) for
  - words within the training set: **0.77**
  - words outside the training set: **0.70**
    - The model needs to extrapolate from words that are distant from those it was trained on
  - discriminating similar words (e.g. “celery” vs. “corn”): **0.62**

frontiers in
HUMAN NEUROSCIENCE
ORIGINAL RESEARCH ARTICLE
published: 23 August 2011
doi: 10.3389/fnhum.2011.00072

Generating text from functional brain images

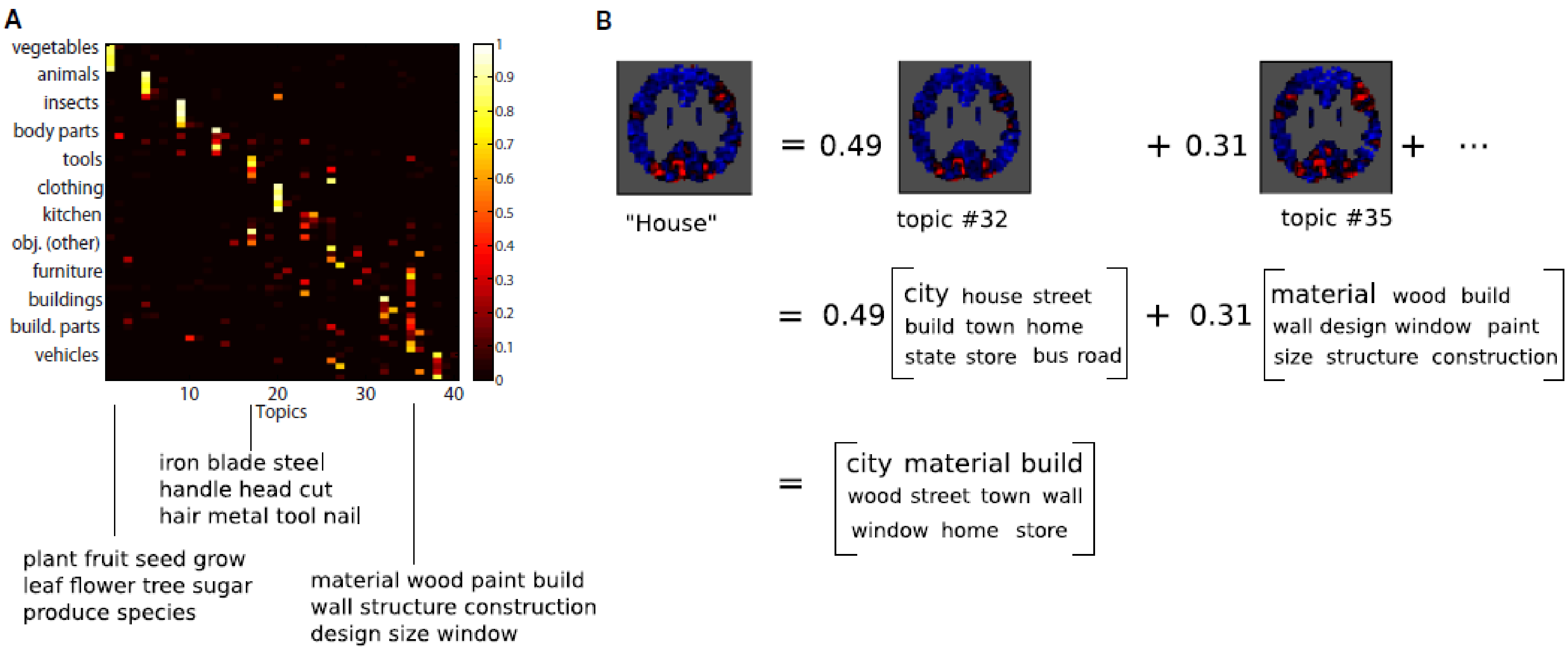
Francisco Pereira[1,2]*, Greg Detre[1,2] and Matthew Botvinick[1,2]

[1] Department of Psychology, Princeton University, Princeton, NJ, USA
[2] Princeton Neuroscience Institute, Princeton University, Princeton NJ, USA



Methodology

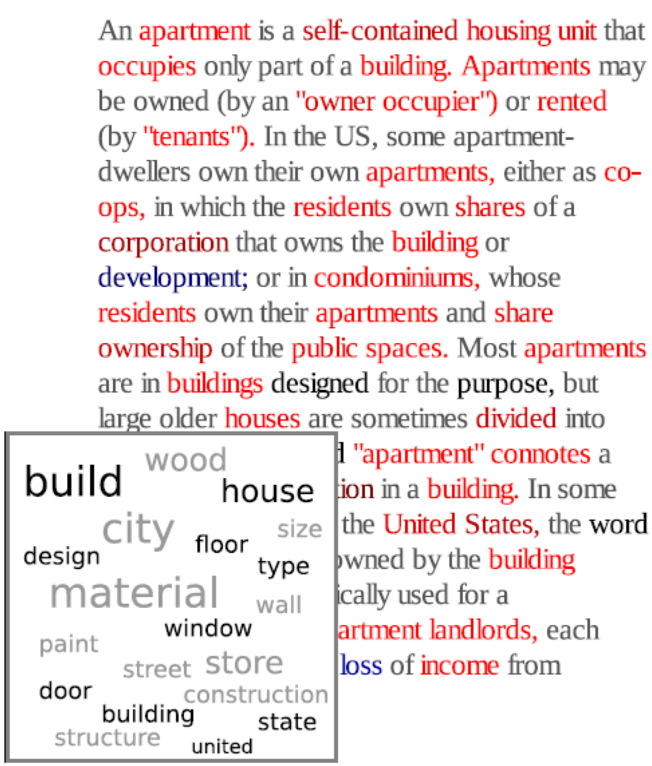
- Corpus of text (Wikipedia articles) used to create a *topic model*, a latent factor representation of each article (the *concept* the article is about).
- Learn a mapping from each latent factor in the model from Step 1 to a corresponding brain image that captures how the factor gives rise to sub-patterns of distributed brain activity, using a training set of brain images.
- For each brain image in a new, test set, the mapping from Step 2 can be used to infer a weighting over latent factors. Given this, *invert* the generative model from Step 1 in order to map from the latent factor representation to *text*, in the shape of a probability distribution over words.



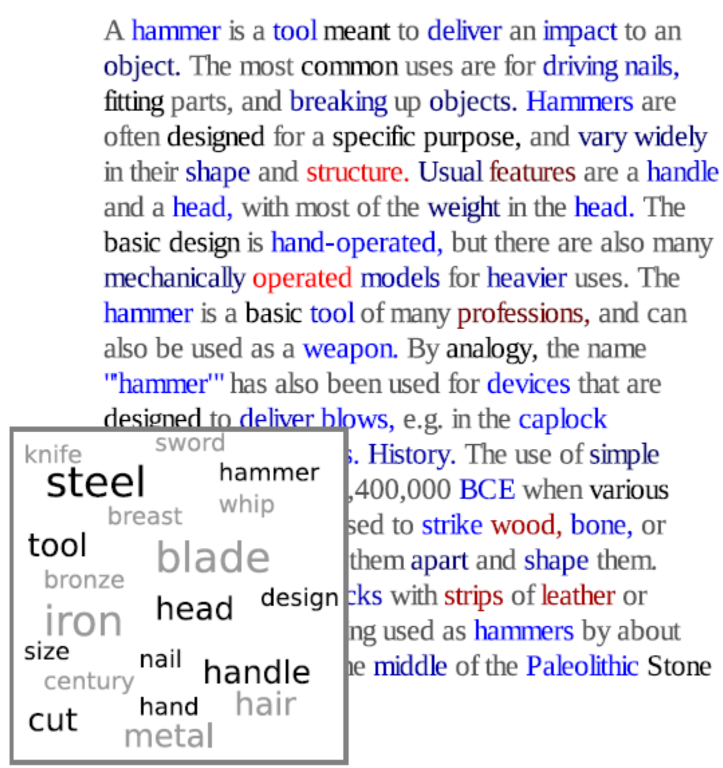
A
vegetables
animals
insects
body parts
tools
clothing
kitchen
obj. (other)
furniture
buildings
build. parts
vehicles
10
20
30
40
Topics
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
iron blade steel
handle head cut
hair metal tool nail
plant fruit seed grow
leaf flower tree sugar
produce species
material wood paint build
wall structure construction
design size window
B
"House"
= 0.49
topic #32
+ 0.31
topic #35
+ ⋯
= 0.49
city house street
build town home
state store bus road
+ 0.31
material wood build
wall design window paint
size structure construction
=
city material build
wood street town wall
window home store

Keywords derived from brain activity of a person reading the article

Apartment



Hammer



Current Biology *21*, 1641–1646, October 11, 2011 DOI 10.1016/j.cub.2011.08.031

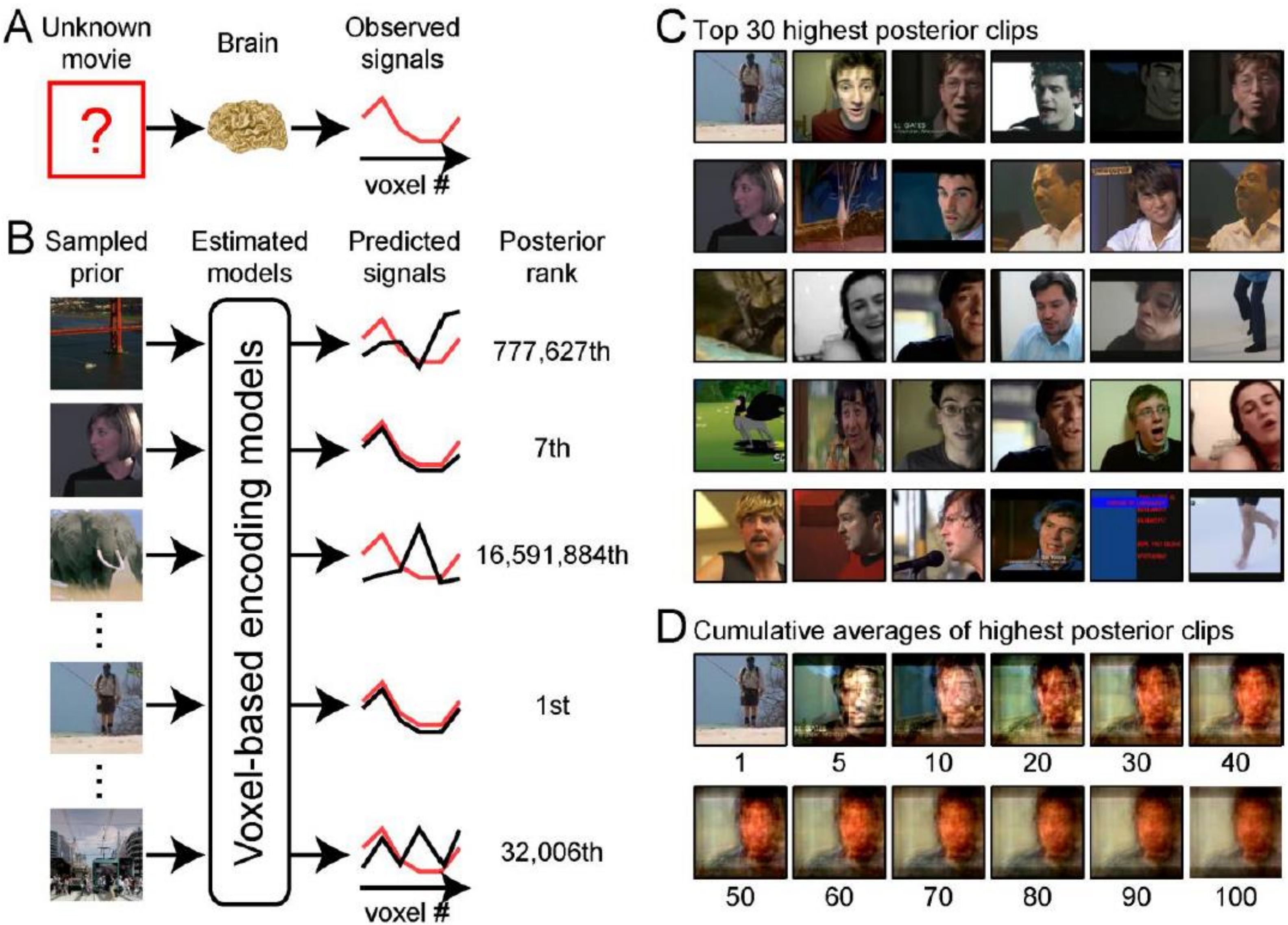
Report

Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies

Shinji Nishimoto,[1] An T. Vu,[2] Thomas Naselaris,[1]
Yuval Benjamini,[3] Bin Yu,[3] and Jack L. Gallant[1,2,4,*]

mental processes. It has therefore been assumed that fMRI data would not be useful for modeling brain activity evoked

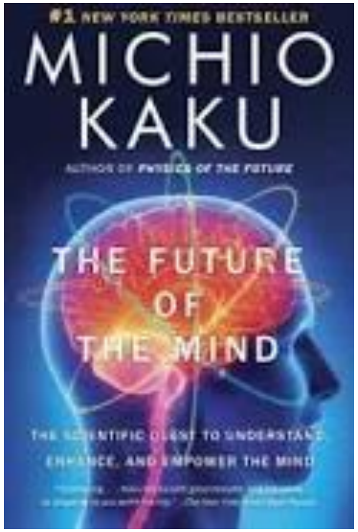
- Record brain activity while the subject watches several hours of movie trailers.
- Build dictionaries (regression model) to translate between the shapes, edges and motion in the movies and measured brain activity. A separate dictionary is constructed for each of several thousand points in the brain at which brain activity was measured.
- Record brain activity to a new set of movie trailers that will be used to test the quality of the dictionaries and reconstructions.
- Build a random library of ~5000 hours of video downloaded at random from YouTube (that have no overlap with the movies subjects saw in the magnet). Put each of these clips through the dictionaries to generate predictions of brain activity. Select the 100 clips whose predicted activity is most similar to the observed brain activity. Average those clips together. This is the reconstruction.


A Unknown movie
Brain
Observed signals
?
voxel #
B Sampled prior
Estimated models
Predicted signals
Posterior rank
Voxel-based encoding models
777,627th
7th
16,591,884th
1st
32,006th
voxel #
C Top 30 highest posterior clips
D Cumulative averages of highest posterior clips
1
5
10
20
30
40
50
60
70
80
90
100

- [video]()

The Future of the Mind (Kaku, 2014)

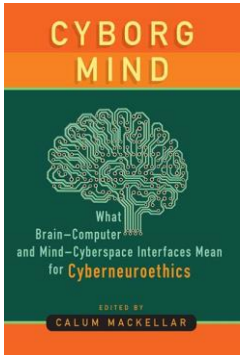
- telepathy
  - speech synthesis from heard word representation
    - ECoG -electrocorticogram - invasive
  - writing from seen images of letters
    - ECoG and EEG
  - nanoelectrodes
- telekinesis
  - prosthetics and exoskeletons
  - videogames
  - paired brains over internet
  - immersive entertainment
  - future
    - brain-net
    - "teleoperated" avatars



The Future of the Mind (Kaku, 2014)

- artificial memories
  - experiments with mice
    - recorded memory in hippocampus, then erased and reactivated
    - false memories
      - optogenetics
    - artificial hippocampus, cerebellum, cortex
    - genetic modifications, drugs
  - forgetting
  - potential for treatment of Alzheimer's disease, PTSD
  - fast learning of new skills
  - vivid "experiences"
  - issues: legal testimony, alibi
- boosting intelligence
  - TMS, drugs, gene therapy (e.g. to simulate eidetic memories or some skills found in people with Asperger's syndrome)
- dream recording

Cyberneuroethics

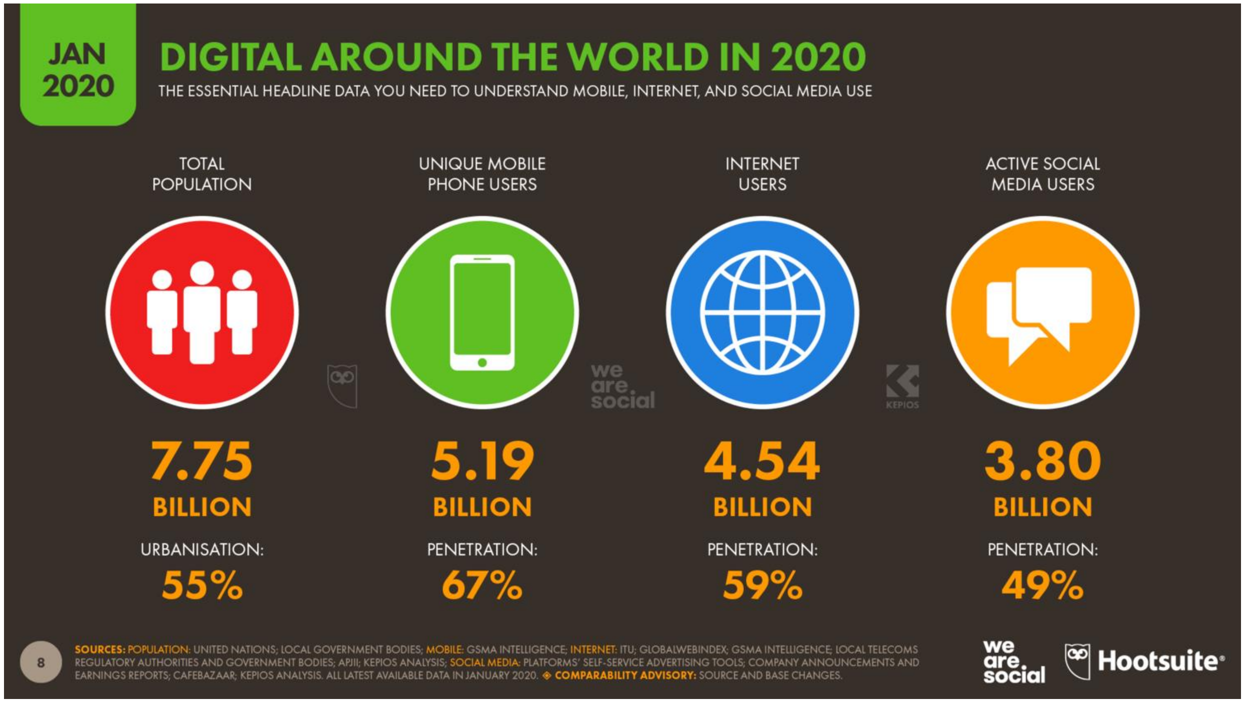


- Individual ethics
  - Advantages: Competitive vs intrinsic benefits
  - Risks: Damage, side-effects, addiction/dependence
- Societal ethics
  - Potential advantages: addressing the limitation of human nature, a fairer society, enable individuals to achieve their full potential.
  - Risks: dependence, social problems and costs, increased pressure/expectations, coercion (device use demanded by society), increased inequality

Issues

- Free will, moral responsibility:
  - Decision making can be affected by education ☺, but also drugs, BCI, …
- Coping with hyper-connectivity

Hyper-connectivity



- internet users worldwide (95% in North America, 87% in Europe, 39% in Africa)
- 6 h : 39 min – average time spent on internet per day by each user

(source: www.internetworldstats.com, 30 May 2020 and DataReportal, April 2020)

Hyper-connectivity

- Inability to cope with the amount of information available
- quickly flicking between many topics may undermine an individual's ability to concentrate
- individuals who are regularly confronted with several streams of electronic information cannot pay attention, control their memory or switch from one job to another as well as those who prefer to complete one task at a time (Stanford, 2009).
- difficult to compartmentalise different parts of their lives, such as work and family life.
- difficult to maintain boundaries between online and offline identities
- Immersivness of virtual worlds and social media
- Cognitive skills can change/adapt

(Not so near) future of minds

Transhumanism

- At the core of transhumanism is the conviction that the **lifespan be extended, aging reversed, and that death should be optional rather than compulsory** (Lee, 2019).
- Means: AI, surgery, genetic engineering, nanotechnology, mind uploading
- Y. Harari: creation of a higher-level species *homo deus:* **permanent happiness, overcoming of aging and diseases, immortality**
- H. Moravec: intelligent machines (*machina sapiens*) will provide us with **personal immortality** and allow us to **evolve independently of human biology and its limitations**

Flying too high?

The Transhumanist Declaration (2002)

1. Humanity stands to be profoundly affected by science and technology in the future. We envision the possibility of broadening human potential by **overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth**.
2. We believe that humanity’s potential is still mostly unrealized. There are possible **scenarios that lead to wonderful and exceedingly worthwhile enhanced human conditions**.
3. We recognize that **humanity faces serious risks, especially from the misuse of new technologies**. There are possible realistic scenarios that lead to the loss of most, or even all, of what we hold valuable. Some of these scenarios are drastic, others are subtle. Although all progress is change, not all change is progress.

The Transhumanist Declaration (2002) – cont.

4. **Research effort needs to be invested** into understanding these prospects. We need to carefully deliberate **how best to reduce risks and expedite beneficial applications**. We also need forums where people can constructively discuss what should be done, and a social order where responsible decisions can be implemented.

5. **Reduction of existential risks, and development of means for the preservation of life and health, the alleviation of grave suffering, and the improvement of human foresight and wisdom should be pursued as urgent priorities, and heavily funded**.

6. **Policy making** ought to be guided by responsible and inclusive moral vision, taking seriously both opportunities and risks, **respecting autonomy and individual rights, and showing solidarity** with and concern for the interests and **dignity of all people** around the globe. We must also consider our **moral responsibilities towards generations that will exist in the future**.

The Transhumanist Declaration (2002) – cont.

7. We advocate the **well-being of all sentience**, including humans, non-human animals, and any future artificial intellects, modified life forms, or other intelligences to which technological and scientific advance may give rise.
8. We favor allowing individuals wide **personal choice** over how they enable their lives. This includes use of techniques that may be developed to **assist memory, concentration, and mental energy; life extension therapies; reproductive choice technologies; cryonics procedures**; and many other possible **human modification and enhancement** technologies.

Transhumanist program

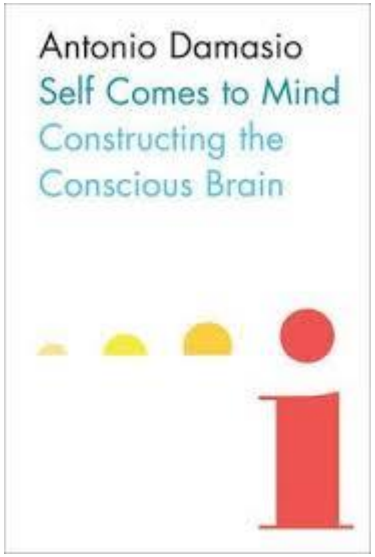
- Enhance human potential
- Eliminate suffering
- Increase life-span
- Achieve immortality

Transhumanist program

- Enhance human potential
- **Eliminate suffering**
- Increase life-span
- Achieve immortality

Mood change toward permanent pleasure

- Regulatory role of affects
- Pain and pleasure evolved to support homeostasis of vital parameters



Suffering as a driving force of progress (Kováč, 2000; 2020)

Escaping reality for cyberspace

- Online identity more ambiguous
- Embodiment vs getting rid of the physical body

Getting rid of the (dependency on) physical body

- Moonshot project from Japan Science and Technology Agency: **[Realization of a society in which human beings can be free from limitations of body, brain, space, and time by 2050]**.
- Outline:
  - To overcome the challenges of a declining birthrate, aging population and associated labor shortage, the key is to realize a society free from the limitations of body, brain, space, and time and allow people with various backgrounds and values – such as the elderly and those with responsibilities for nursing and childcare – to actively participate in society.
  - Our R&D will develop core technologies related to cyborgs and avatars, called ‘Cybernetic Avatars’, allowing expansion of human physical, cognitive and perceptual abilities. We will build ‘Cybernetic Avatar Infrastructure’ in the cloud while easing the acceptance of Cybernetic Avatars into future society.

Transhumanist program

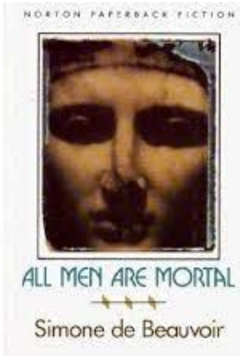
- Enhance human potential
- Eliminate suffering
- **Increase life-span**
- Achieve immortality

Societal disruption by increased longevity

- Human enhancement has critical implications for **how society is organized**. Healthier people will mean the prospect of **longer lives**, which in turn will mean a growing ageing population.
- Impact on social provisions and the broader economic infrastructure of a society, requiring people and governments to revise their expectations about the **duration of the working life**, the **economics of pension funds**, and the **provision of health insurance**.
- Demands on social systems that may bring about their collapse, if they are not rethought.

Living long life vs meaningful life

- “The mystery of human existence lies not in just staying alive, but in finding something to live for.” (F.M. Dostoyevsky in *The Brothers Karamazov*)
- “It is not that we have a short time to live, but that we waste a lot of it. Life is long enough, and a sufficiently generous amount has been given to us for the highest achievements if it were all well invested.” (Seneca)
- If you had one day left to live, how would you spend it?
- If you are to live a thousand more years, what would you do with all that time?

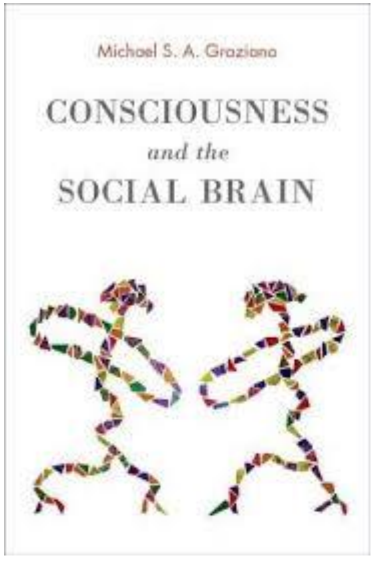


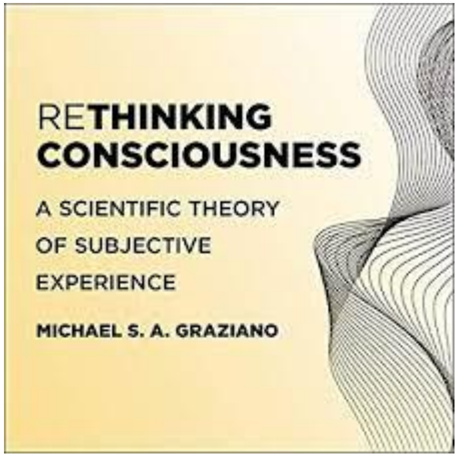
Transhumanist program

- Enhance human potential
- Eliminate suffering
- Increase life-span
- **Achieve immortality**

Mind uploading

- Gradual replacement
- Destructive scan and copy





Problems (MacKellar, 2019)

- Distinction between reality and virtuality would disappear. Cyber-individuals could be trapped in an existence in which virtual dangers and nightmares become as real as their own reality.
- Since backups of a person would need to be created to protect against viral attack or sudden catastrophic failure of a main drive, how can an individual be sure that these are safe and secure?
- Could the backups, themselves, be considered as persons brought into existence through a copying procedure of the original person?
- Who has access to these backups? Breach of security would be cyberspace's equivalent to a forced entry or personal trespass. Hacking would be a personal invasion on a level perhaps in the same category as other violent invasive crimes such as rape.
- What would happen if the backups and the files expressing a person were irreversibly lost? Would this then represent a form of death of the individual?
- How would concepts such as compassion and empathy which make existences meaningful, be able to be expressed in digital persons, since these notions require a capacity to suffer?

Death (population replacement) as the driving force of evolution

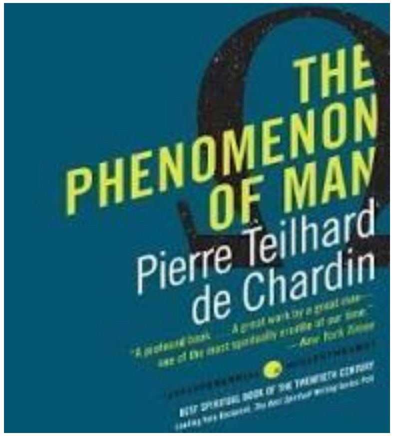
- Gerontocracy
- Importance of finiteness of life for human value system

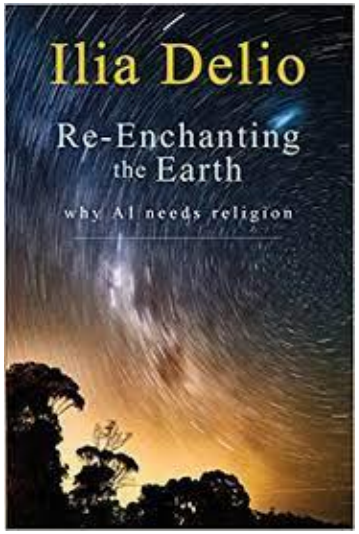
Criticism of transhumanism as individual betterment

- Participation, not self-fulfillment, a transcendence towards something more, not perfection of a limited existence.
- “Technologies of the self, whether a cyber self or a new genetic self are self-asserting rather than self-transforming, enhancing the ego rather than surrendering it to a greater reality and purpose.” (R. Cole Turner in Delio, 2011).

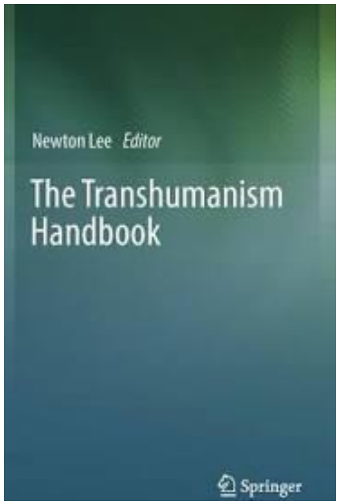
Network intelligence

- Hyper-connected mind / collective consciousness
  - Kurzweil (2005) predicted that Singularity would happen in 2045
  - Pierre Teilhard de Chardin (1940s) predicted internet and planetarisation of consciousness (noosphere)
- Transhumanism and religion

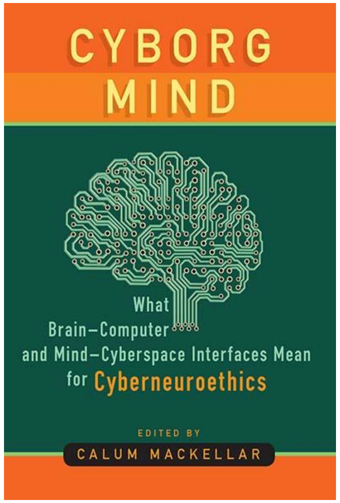




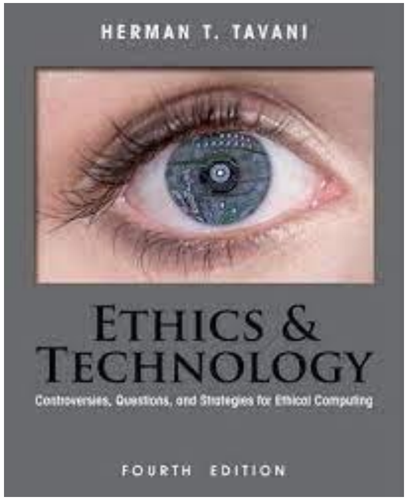
References



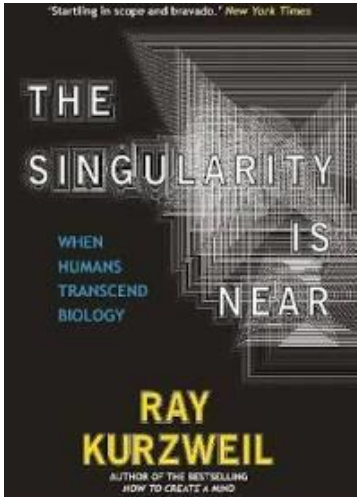
Newton Lee Editor
The Transhumanism Handbook
Springer



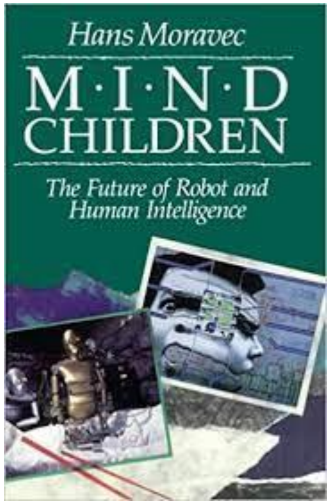
CYBORG MIND
What Brain–Computer and Mind–Cyberspace Interfaces Mean for Cyberneuroethics
EDITED BY
CALUM MACKELLAR



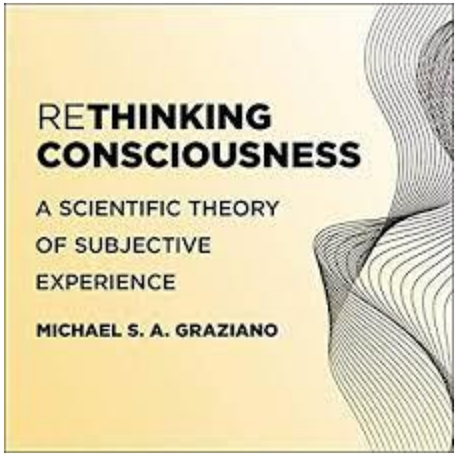
HERMAN T. TAVANI
ETHICS & TECHNOLOGY
Controversies, Questions, and Strategies for Ethical Computing
FOURTH EDITION



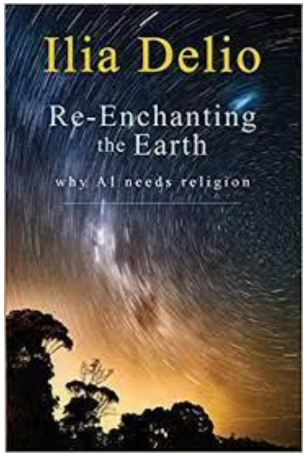
'Startling in scope and bravado.' New York Times
THE SINGULARITY IS NEAR
WHEN HUMANS TRANSCEND BIOLOGY
RAY KURZWEIL
AUTHOR OF THE BESTSELLING HOW TO CREATE A MIND



Hans Moravec
M·I·N·D CHILDREN
The Future of Robot and Human Intelligence



RETHINKING CONSCIOUSNESS
A SCIENTIFIC THEORY OF SUBJECTIVE EXPERIENCE
MICHAEL S. A. GRAZIANO



Ilia Delio
Re-Enchanting the Earth
why AI needs religion