Assessing the impact of AI on human behaviour: interdisciplinary views

Emilia Gómez
@emiliagogu

Joint work/slides from Vicky Charisi, Marius Miron, Songül Tolan, Nando Martínez-Plumed, Enrique Fernández-Macías, Annarosa Pesole, Maria Iglesias (JRC); Carlos Castillo, Lorenzo Porcaro (UPF); José H. Orallo (UPV); Luis Merino, Fernando Caballero (UPO); Bob Sturm (KTH); Emilio Gómez-González (US)
Outline

- Motivation
- Interdisciplinarity and diversity
- Selected projects
Outline

- Motivation
- Interdisciplinarity and diversity
- Selected projects

Assessing the impact of AI on human behaviour
Hatsune Miku

https://www.youtube.com/watch?v=dhYaX01NOfA
Kondo "marries" a moving, talking hologram

https://www.youtube.com/watch?v=dtu4t_Zc3d4

Artificial Intelligence

Machines or agents capable of observing its environment and taking decisions towards a certain goal

- Machine learning: data+computation+algorithms
  - General purpose (GPT)
  - Scalable, personalization
  - Address cognitive tasks
Deep learning method, data-driven

Stanford algorithm can diagnose pneumonia better than radiologists

Stanford researchers have developed a deep learning algorithm that evaluates chest X-rays for signs of disease. In just over a month of development, their algorithm outperformed expert radiologists at diagnosing pneumonia.

By Taylor Knasko
Stanford researchers have developed an algorithm that offers diagnoses based on chest X-ray images. It can diagnose up to 14 types of medical conditions and is able to diagnose pneumonia better than expert radiologists working alone. A paper about the algorithm, called CheXNet, was published Nov. 14 on the open-access, scientific preprint website arXiv.

"Interpreting X-ray images to diagnose pathologies like pneumonia is very challenging, and we know that there's a lot of variability in the diagnoses radiologists arrive at," said Prateek Rajpurkar, a graduate student in the Stanford Machine Learning Group and co-author of the paper. "We became interested in developing machine learning algorithms that could have fast, highly accurate predictions and make accurate.

"When the head of the U.S. Supreme Court said artificial intelligence (AI) is having a significant impact on how the legal system thinks about issues, we pay attention. That's exactly what Judge Andrea John found in her research," said John.

"The idea was to create a system that could help doctors and patients make better decisions."
Technology impact assessment

1. Who are the people affected?
2. Who are the ‘winners’ (benefit), who the ‘losers’ (cost)?
3. How many lives can be saved?
4. How much money/jobs can be saved?
5. What are the short-term and long-term costs/benefits?
Human behaviour and machine intelligence

Provide cognitive assistance: “computer-assisted”.

Affect decision making and cognitive and socio-emotional capabilities.

GOALS
Find the right balance
Best strategies for human-AI competition cooperation == human-centered AI
HUMAINT key research principles

1. Interdisciplinary

- Human behaviour
- Machine learning
- Economics of AI

JRC
HUMAINT key research principles

1. Interdisciplinary
2. Impact

We publish open, reproducible research
We provide policy support
HUMAINT key research principles

1. Interdisciplinary
2. Impact
3. Community

- 8 associated external fellows
- Universidad Pablo Olavide
- Universidad de Sevilla
- Universidad Politécnica de Valencia
- University of Cambridge
- International Consortium for Socially Intelligent Robotics
- TROMPA (*Towards Richer Online Music Public-domain Archives*) H2020 project
- Other JRC units and DGs of the EC
- AIST Japan, Honda Research Institute

**Human behaviour**

**Research community**

**Machine learning**

**Economics of AI**
Outline

- Motivation
- Interdisciplinarity and diversity
- Selected projects

This summary is based on the work by Barry, A., Born, G., and Weszkalnys, G. Logics of interdisciplinarity. Economy and Society Volume 37 Number 1 February 2008:20-49
Disciplines discipline disciples

A commitment to a discipline is a way of ensuring that certain disciplinary methods and concepts are used rigorously and that undisciplined and undisciplinary objects, methods and concepts are ruled out.
Why interdisciplinarity?

1. Accountability

1. Innovation and economic growth

2. Ontology: affect ontological change.

What is your main discipline?

Copy and edit your own interdisciplinarity sheet
https://tinyurl.com/tuknmd5
Beyond disciplines

- Boundary transgressions
- Solution to a series of contemporary problems.
- New model of knowledge production: new forms of quality control (Nowotny, Scott and Gibbons, 2001)


https://www.flickr.com/photos/frauleinschiller/5612922237
Can you identify several disciplines in your work?

[Image of a mind map showing various disciplines including Social-sciences, Natural sciences, Arts, and Humanities]

https://upload.wikimedia.org/wikipedia/commons/7/7c/Disciplines_mind_map.jpg
My main disciplines

Disciplines:
- Musicology
- Information retrieval
- Ethnomusicology
- Music cognition

Professions:
- Physical performance
- Recreation
- Environmental studies
- Forestry
- Agriculture
- Medicine
- Family-Consumer science
- Social work
- Public administration
- Law
- Human history
- Religion
- Literature
- Journalism Media Communication
- Philosophy
- Linguistics
- Library Museum studies
- Intelligence
- Military sciences
- Divinity
- Engineering Technology
- Architecture Design
- Transportation
- Business

Humanities:

Social sciences:
- Archaeology
- Area studies
- Geography
- Anthropology
- Cultural Ethnic studies
- Gender Sexuality studies
- Psychology
- Sociology
- Economics
- Political science

Natural sciences:
- Physics
- Chemistry
- Biology
- Earth sciences
- Space sciences

Formal sciences:
- Computer sciences
- Logic
- Mathematics
- Pure mathematics
- Statistics
- Systems science

Arts:
- Journalism Media Communication
- Philosophy
- Linguistics

ISMIR

https://upload.wikimedia.org/wikipedia/commons/7/7c/Disciplines_mind_map.jpg
Modes and logics of interdisciplinarity

Interdisciplinarity is not historically novel BUT there is a new sense that it is a need to better connect research & society/economy.

Methodology (Barry, Born and Weszkalnys, 2008)

- Internet-based mapping survey of interdisciplinary fields.
- Selected fields:
  a. Environmental and climate change research
  b. Ethnography in the IT industry
  c. Art-science
- 10 case studies of initiatives in these fields across different national settings.
Concepts

1. Multidisciplinarity
   ○ Several disciplines cooperate but remain unchanged, working with standard disciplinary framings.

2. Interdisciplinarity
   ○ Integrate or synthesize perspectives from several disciplines.

3. Transdisciplinary
   ○ Transgression, fusion.
   ○ Oriented to the complexity of real-world problem solving, overcoming distance between specialized and lay knowledges or between research and policy.

Modes of interdisciplinarity

1. **Integrative-synthesis**: integration of disciplines in relatively symmetrical form.
   - Example: synthesis of disciplines via “universal” mathematical models: climate change research integrating natural scientific and social scientific accounts for impact.

1. **Subordination-service**: master vs service discipline.
   - Example: art to communicate science, science as a service to art (providing resources and equipment for a project conceived in artistic term).

1. **Agonistic-antagonistic**: criticism to transcend historical disciplines into new ones.
   - Example: ethnography in the IT industry as an opposition to previous sociological approaches to the study of technology or to scientific approaches to study technologies.
Modes of interdisciplinarity & methodologies

1. Integrative-synthesis

1. Subordination-service

1. Agonistic-antagonistic

Methodological orientations

- Problem-solving, policy orientation in response to new problems/objects.
- Practice-oriented, labour division.

Can you identify modes of interdisciplinarity or methodologies?

https://upload.wikimedia.org/wikipedia/commons/7/7c/Disciplines_mind_map.jpg
Diversity

- Interdisciplinarity is a particular aspect of diversity.
- Valued for incorporating different views in the design process.
- Diversity is difficult to conceptualize.
  - Disciplines
  - Cultural background
  - Gender
  - ...

Fostering diversity, AI systems should be accessible to all, regardless of any disability, and involve relevant stakeholders throughout their entire life circle.

Figure 1: Schematic representation of the attributes of diversity, in the context of interdisciplinary analysis, from (Rafols and Meyer 2010).

divinAI (divinAI.org)

- Collaborative project: Universitat Pompeu Fabra, Joint Research Centre, welcoming contributors
- Study how diverse are AI conference, related to AI geo-politics
- Define a set of indicators derived from biodiversity (Pielou, Shannon Index).
  - Gender
  - Geographical origin, institution (culture)
  - Focus (academia vs industry)
- Monitor the distribution, evolution, impact of diversity policies.
- Hackfest Barcelona 31st, New York February 10th

ICML 2017

divinai.org
Universitat Pompeu Fabra, Barcelona, Jan 31st
AAAI diversity & inclusion activities, New York
Take-home messages

● Benefits of interdisciplinary approaches to address societal problems.
● Interdisciplinarity takes many forms.
● Not easy to achieve transdisciplinarity: vocabulary, methods, quality standards.
● Has practical risks.
● Link with diversity of communities.
Policy making questions

1. How can AI affect human decision making? e.g. recidivism prediction
2. How does social robots affect children development?
3. How will AI impact jobs and workplaces?
4. Which dual use can have AI in medicine/healthcare?
5. How will recommender systems impact opinion/culture?
Outline

- Motivation
- Interdisciplinarity and diversity
- Selected projects
HUMAINT research topics

1. Decision making
2. Child-robot interaction
3. AI and EU labour markets
4. Medicine and healthcare
5. Music
HUMAINT research topics

1. Decision making
2. Child-robot interaction
3. AI and EU labour markets
4. Medicine and healthcare
5. Music

Decision making: humans

- Humans are prone to cognitive biases (Kahneman, 2011)
- Judge decisions can be affected by hunger or mood (Dazinger et al., 2011; Chen et al., 2016)

Bias, fairness, discrimination

- **Bias**: systematic deviation from truth, a feature of statistical models (Metcalf, 201).
- **Fairness**: a feature of value judgments (Metcalf, 2019)
- **Discrimination**: a legal concept based on group membership:
  - sex, race, colour, ethnic and social origin, political opinion, membership of a national minority, property, birth, disability, age or sexual orientation
  (Article 14, European Convention on Human Rights)
Decision making: **ML algorithms**

- Support the formalization of decision making process
- Not neutral, may learn human biases (Barocas and Selbs, 2016; Angwin et al., 2016)
- Reliance, liability & responsibility

Decision making: humans vs algorithms

1. Data on human decision making
2. Model, evaluate and understand: predictive performance and group fairness* (human and interpretable ML models)
3. Design best cooperation strategies

- Task: binary classification

* Computer science researchers talk of at least 21 definitions of fairness
Decision making: humans vs algorithms

1. Data on human decision making
2. Model, evaluate and understand: predictive performance and group fairness* (human and interpretable ML models)
3. Design best cooperation strategies

- Task: binary classification
- SAVRY: structured professional risk assessment framework (24 risk factors, final assessment)
- Antonio Pueyo (Universitat de Barcelona), Carlos Castillo (Universitat Pompeu Fabra)

Decision making: **humans vs algorithms**

- Machine Learning improves predictive performance
- BUT may lead to unfairness…
- Algorithms emphasize correlations (base rates)

Static features: defendant demographics and criminal history
SAVRY scores: expert assessment (24)

ML: logistic regression (logit), multi-layer perceptron (mlp), support vector machine with a linear (lsvm) or radial (rsvm) kernel, K-nearest neighbors (knn), random forest (rf), and naive bayes (nb)

---

Decision making: **humans vs algorithms**

- Machine Learning improves performance
- BUT may lead to unfairness…
- Algorithms emphasize correlations (base rates)

Static features: defendant demographics and criminal history  
SAVRY scores: expert assessment (24)

ML: logistic regression (logit), multi-layer perceptron (mlp), support vector machine with a linear (lsvm) or radial (rsvm) kernel, K-nearest neighbors (knn), random forest (rf), and naive bayes (nb)

---

Algorithm-supported decision making

- Data limited, specially in sensitive and complex scenarios.
- Developers must understand the social context in which the algorithm will be embedded (Selbst et al. 2019).
- Domain experts and users must understand the algorithmic approach (transparency).
# Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Mode of interdisciplinarity</th>
<th>Methodological orientations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Integrative synthesis</td>
<td>Problem-solving</td>
</tr>
<tr>
<td></td>
<td>Subordination service</td>
<td>Practice-oriented</td>
</tr>
<tr>
<td></td>
<td>Agonistic antagonistic</td>
<td>Other</td>
</tr>
</tbody>
</table>
Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Other disciplines</th>
<th>Computer Science, mathematics (formal sciences)</th>
<th>Psychology (social sciences)</th>
<th>Sociology (social sciences)</th>
<th>Law (humanities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of interdisciplinarity</td>
<td>Integrative synthesis</td>
<td>Subordination service</td>
<td>Agonistic antagonistic</td>
<td>Fair Machine Learning</td>
</tr>
<tr>
<td>Methodological orientations</td>
<td>Problem-solving</td>
<td>Practice-oriented</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluation</td>
<td>Recidivism prediction</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HUMAINT research topics

1. Decision making
2. Child-robot interaction
3. AI and EU labour markets
4. Medicine and healthcare
5. Music

Child-Robot Interaction

- Social robots = embodied AI

Human-robot interaction: disciplines

Disciplines
- Embodied Social AI
- AI (Machine learning)
- Robotics
- Psychology
- Philosophy

Child-robot interaction: approach

1. Behavioural studies: problem solving, social interaction, emotional engagement
2. Qualitative & quantitative understanding
3. Experiment and design interaction strategies and contribute to system design

- Task: tower of Hanoi
- 5-8 y.o. children
- Social robot

Lucas, 1883

Gómez, R. Haru: Hardware Design of an Experimental Tabletop Robot Assistant, HRI2018
Child-robot interaction: approach

- Study 1
- What is the impact of Social Robot Interventions on children’s learning process?

Child-robot interaction: approach

Methodology
- 72 sessions of 15 min, 113 tasks from 20 children.

Results

- Need for exploration
- Importance of self-initiated interaction
- Individual differences
- Learning process

# Ethical considerations with children

<table>
<thead>
<tr>
<th>Research ethics</th>
<th>Responsible design and innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What</strong></td>
<td></td>
</tr>
<tr>
<td>Transparency</td>
<td>- How does research affect the paradigms in <strong>formal education and informal learning</strong>?</td>
</tr>
<tr>
<td>Privacy</td>
<td>- How will <strong>industry</strong> of Children’s Toys and Media be aligned with the Child’s Values and Rights?</td>
</tr>
<tr>
<td>Consistency</td>
<td>- How can we <strong>embed Child’s Values and Rights</strong> into our systems?</td>
</tr>
<tr>
<td>Explainability</td>
<td></td>
</tr>
<tr>
<td>Inclusion</td>
<td></td>
</tr>
<tr>
<td>Deception</td>
<td></td>
</tr>
</tbody>
</table>
Ethical considerations: what

Priorities
- From principles and policies to practice
- Clearer concepts and more evidence
- Children’s agency and data
- Broad stakeholder agency

Ethical considerations: how

Society does not have universal standards or guidelines to help embed human norms or moral values into autonomous intelligent systems (AIS) today. But as these systems grow to have increasing autonomy to make decisions and manipulate their environment, it is essential they be designed to adopt, learn, and follow the norms and values of the community they serve, and to communicate and explain their actions in as transparent and trustworthy manner possible, given the scenarios in which they function and the humans who use them.

The conceptual complexities surrounding what "values" are make it currently difficult to

https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/eadie_embedding_values.pdf
# Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Mode of interdisciplinarity</th>
<th>Methodological orientations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Integrative synthesis</td>
<td>Problem-solving</td>
</tr>
<tr>
<td></td>
<td>Subordination service</td>
<td>Practice-oriented</td>
</tr>
<tr>
<td></td>
<td>Agonistic antagonistic</td>
<td>Other</td>
</tr>
</tbody>
</table>
Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Psychology (social sciences)</th>
<th>Engineering and technology (applied science)</th>
<th>Computer Science, mathematics (formal sciences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of interdisciplinarity</td>
<td>Integrative synthesis</td>
<td>Subordination service</td>
<td>Agonistic antagonistic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Technology as a tool for human-AI interaction</td>
<td>Human-Robot Interaction</td>
</tr>
<tr>
<td>Methodological orientations</td>
<td>Problem-solving</td>
<td>Practice-oriented</td>
<td>Other</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Design human-robot interactions</td>
<td></td>
</tr>
</tbody>
</table>
HUMAINT research topics

1. Decision making
2. Child-robot interaction
3. AI and EU labour markets
4. Medicine and healthcare
5. Music

Humans carry out tasks at work

<table>
<thead>
<tr>
<th>Content</th>
<th>Methods and tools</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Physical tasks</strong></td>
<td><strong>1. Work organisation</strong></td>
</tr>
<tr>
<td>(a) Strength</td>
<td>(a) Autonomy</td>
</tr>
<tr>
<td>(b) Dexterity</td>
<td>(b) Teamwork</td>
</tr>
<tr>
<td><strong>2. Intellectual tasks</strong></td>
<td>(c) Routine</td>
</tr>
<tr>
<td>(a) Information processing:</td>
<td>(I) Repetitiveness</td>
</tr>
<tr>
<td>(I) L.P. of uncodified information</td>
<td>(II) Standardization</td>
</tr>
<tr>
<td>(II) L.P. of codified information</td>
<td></td>
</tr>
<tr>
<td>(i) Literacy:</td>
<td></td>
</tr>
<tr>
<td>(a) Business</td>
<td></td>
</tr>
<tr>
<td>(b) Technical</td>
<td></td>
</tr>
<tr>
<td>(c) Humanities</td>
<td></td>
</tr>
<tr>
<td>(ii) Numeracy:</td>
<td></td>
</tr>
<tr>
<td>(a) Accounting</td>
<td></td>
</tr>
<tr>
<td>(b) Analytic</td>
<td></td>
</tr>
<tr>
<td>(b) Problem solving:</td>
<td></td>
</tr>
<tr>
<td>(I) Information gathering and evaluation</td>
<td></td>
</tr>
<tr>
<td>(II) Creativity and resolution.</td>
<td></td>
</tr>
<tr>
<td><strong>3. Social tasks</strong></td>
<td></td>
</tr>
<tr>
<td>(a) Serving/attending</td>
<td></td>
</tr>
<tr>
<td>(b) Teaching/training/coaching</td>
<td></td>
</tr>
<tr>
<td>(c) Selling/influencing</td>
<td></td>
</tr>
<tr>
<td>(d) Managing/coordinating</td>
<td></td>
</tr>
</tbody>
</table>

Source: Fernández-Macías and Bisello [2017]
Machine intelligence impact

- Technology increases the productivity of all workers, particularly high-skilled workers (Katz and Murphy, 1992)
- Technology also performs labour substitution, polarization
- Approach: task-based framework + work organization (Autor, 2014a,b, Autor et al., 2003; Acemoglu and Autor, 2011)
- We focus on Machine Learning techniques
- We use cognitive abilities as an intermediate step (Hernández-Orallo, 2017)

**Cognitive abilities:**

- Memory processes
- Sensorimotor interaction
- Visual processing
- Auditory processing
- Attention and search
- Planning and sequential decision making and acting
- Comprehension and compositional expression
- Communication
- Emotion and self-control
- Navigation
- Conceptualisation, learning and abstraction
- Quantitative and logical reasoning
- Mind modeling and social interaction
- Metacognition
From labour to ML paradigms

From labour to ML paradigms

- Delphy method
- Several rounds of questionnaires to experts
- People do tasks differently than machines
- Discussion and refinement

### Table 1: Difference in annotations between round 1 and round 2

<table>
<thead>
<tr>
<th></th>
<th>abilities</th>
<th>consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>round 1</td>
<td>round 2</td>
</tr>
<tr>
<td>Average</td>
<td>6.03</td>
<td>5.34</td>
</tr>
<tr>
<td>Min</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Max</td>
<td>13</td>
<td>10</td>
</tr>
</tbody>
</table>

From labour to ML paradigms

- Analysis, evaluation, comparison and classification of AI systems.
- Data gathered from scientific papers, experiments, benchmarking initiatives.

Preliminary conclusions

- ML development has mainly addressed perceptual tasks, e.g. visual and auditory perception
- High percentage of tasks assisted by AI
- AI paradigms towards information processing, memory
- AI benchmarking addressing social skills


Work organization

- More than a sum of tasks.
- Generality, autonomy, sociability.
- Work organization.
- Digital labour platforms (e.g. Uber, Amazon Mechanical Turk, Task Rabbit): discrete and granular tasks, algorithmically centralised decision making, standardise processes and outputs.

### Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Mode of interdisciplinarity</th>
<th>Methodological orientations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Integrative synthesis</td>
<td>Problem-solving</td>
</tr>
<tr>
<td></td>
<td>Subordination service</td>
<td>Practice-oriented</td>
</tr>
<tr>
<td></td>
<td>Agonistic antagonistic</td>
<td>Other</td>
</tr>
</tbody>
</table>
Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Economics (social sciences)</th>
<th>Computer Science, mathematics (formal sciences)</th>
<th>Sociology (social sciences)</th>
<th>Psychology (social sciences)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Mode of interdisciplinarity</th>
<th>Integrative synthesis</th>
<th>Subordination service</th>
<th>Agonistic antagonistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematical model</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methodological orientations</th>
<th>Problem-solving</th>
<th>Practice-oriented</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantification</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
HUMAINT research topics

1. Decision making
2. Child-robot interaction
3. AI and EU labour markets
4. **Medicine and healthcare**
5. Music
AI in medicine and healthcare

- Clinical decision making

Machine learning in clinical decision making

DATA ACQUISITION
Best raw data - smallest learning curve

FEATURE EXTRACTION
Best measurements - no tedious work

INTERPRETATION
Probability on normality or differential diagnosis

DECISION SUPPORT
Provide personalised suggestions and prognosis

Machine learning in clinical decision making

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>May enhance the prediction of clinical outcomes</td>
<td>Need well-curated, representative databases for training</td>
</tr>
<tr>
<td>May enhance the prediction of response to treatment</td>
<td>Affected by data reliability, representativeness, competenseness, and bias</td>
</tr>
<tr>
<td>May improve the recommendation of interventions</td>
<td>Need to prove clinical benefit</td>
</tr>
<tr>
<td></td>
<td>Need to be explainable rather than interpretable</td>
</tr>
<tr>
<td></td>
<td>Need to be integrated within clinical systems</td>
</tr>
<tr>
<td></td>
<td>Need to prove cost-effectiveness</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower cost of healthcare by suggesting cost-effective decisions</td>
<td>Harm patients if wrong decisions are taken – high-risk</td>
</tr>
<tr>
<td></td>
<td>Make decisions for the average patient, not at the individual level</td>
</tr>
</tbody>
</table>

General challenges on ML for clinical decision making

- **Learning**
  - Non-standardized data
  - Bias and confounding
  - Continuous validation

- **Accountability/traceability**
  - Interpretability (slow reasoning) vs explainability (Deep Learning): main limiting factors for adoption.
  - Casual ML rather than predictive ML

- **System-related**
  - Security
  - Regulatory
  - Human-machine interaction
  - Real clinical data

Beyond clinical decision making

- Literature review of 582 publications, product descriptions, medical perspective
  - Clinical decision-making
    - Radiology, surgery with augmented reality and surgical robots
    - Followed by other image-based specialties (e.g. pathology, dermatology, ophthalmology)
    - Virtually all areas, from general practitioners to emergency departments, epidemiology, and disease management
  - Online assistants (e-doctors), clinical companions
  - Wearables and IoT → real-time monitoring
  - Genetic tests in an affordable way
Classification

TAL 0. Unknown status. Not considered feasible according to references.
TAL 1. Unknown status. Considered feasible according to related, indirect references.
TAL 2. General/basic idea publicly proposed.
TAL 3. Calls for public funding of R&D open.
TAL 4. Results of academic/partial projects disclosed.
TAL 5. Early design of product disclosed.
TAL 6. Operational prototype/‘first case’ disclosed.
TAL 7. Products disclosed but not available.
TAL 8. Available for restricted (e.g. professional) users.

Algorithms for computer-aided diagnosis. SW for decision support in (most) clinical areas, 8, 9
Structured reports, eHealth. SW for improved workflows, efficiency, 8, 9
AR/VR, advanced imaging tools. Tools for information visualization and navigation, 6, 7, 9
Image-guided surgery, Teleoperation, 6, 6, 9
Digital pathology, ‘viroscopy’. SW for automated, extensive analysis, 4-9
Personalized, precision medicine. Tailored treatments, Prediction of response, 4-9
‘In-silico’ modelling and testing. The ‘digital twin’, 4-9
Drug design, 4, 8
Apps, chatbots, dashboards, online platforms. The ‘digital doctor’ (assistance for professionals and for patients), 8, 9
Big Data collection and analysis. Epidemiology, prevention and monitoring of disease outbreaks, 2-9
IoT, wearables, mHealth. Automated clinical/health surveillance in any environment/institution, 7, 8
Genetic testing, Population screening. Disease tests, Direct-to-consumer tests, 4-9
Personalized, precision medicine. Individual profiling, Personalized molecules for treatment at ‘impossible’ prices, 3-8
Gene editing. ‘Engineered’ humans, 2, 6
Brain-machine interfaces, 5-8
Control of prostheses, exoskeletons, ‘Cyborgs’, 2-7
Neurostimulation, Neuromodulation, 4-8
Neuroprostheses (for the central nervous system), 2-5
Mind-reading and ‘manipulation’, 1-3
Reading and decoding brain signals. Interaction with neural processes, 6, 6
Tailored marketing (e.g. related to female cycles), 6, 6

Ethical and social impact

1. Currently under analysis
2. Of particular relevance in this context
3. Barely addressed, specific

Challenges:
- Extended personalized medicine
- Doctor replacement/enhancement → patient-centred view
- Affordability / inequalities
- Dual use of technology

# Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Mode of interdisciplinarity</th>
<th>Methodological orientations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Integrative synthesis</td>
<td>Problem-solving</td>
</tr>
<tr>
<td></td>
<td>Subordination service</td>
<td>Practice-oriented</td>
</tr>
<tr>
<td></td>
<td>Agonistic antagonistic</td>
<td>Other</td>
</tr>
</tbody>
</table>
# Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Engineering and Technology (applied sciences)</th>
<th>Medicine and health (applied sciences)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mode of interdisciplinarity</strong></td>
<td>Integrative synthesis</td>
<td>Subordination service</td>
</tr>
<tr>
<td><strong>Methodological orientations</strong></td>
<td>Problem-solving</td>
<td>Practice-oriented</td>
</tr>
</tbody>
</table>
HUMAINT research topics

1. Decision making
2. Child-robot interaction
3. AI and EU labour markets
4. **AI and music**
AI also impacts music

- Exploited in all stages, from creation to distribution (platforms)
- Various participants contributing to and benefiting from music: composers, musicians, educators, listeners, and organisations.
- Focus on 2 contexts
  a. AI for music creation: realistic synthesis/composition
  b. AI for music recommendation

Taryn Southern 2017
Impact of AI on music creativity

- Collaboration with Bob Sturm (Computer science), María Iglesias (Law). Oded Ben-Tal (Music composer).
- Copyright law & Engineering practice
- Around folk-NN project https://folkrnn.org/ generate a folk tune with a recurrent neural network. https://www.youtube.com/watch?v=EC1TrQz

Deep bach https://www.youtube.com/watch?v=QiBM7-5hA6o
AI for music creativity: questions

1. In many areas technology leads to more efficient production lines and increased profit but human redundancy and deskillling. Can the same happen in music?
2. Who (and how) is accountable for music-AI systems?
3. Who owns the rights to the music generated by AI models? What is their artistic value?
4. Should musicians be informed about the involvement of AI in the music they play, much the same way ingredients of food products are communicated? What about composers using AI tools?
5. How should this information be presented in a transparent way, and to what level of detail?

Some findings

Copyright Law perspective

- Authorship recognition & copyright may require an analysis of the operation of the systems and the role of the different actors involved (e.g. developer, trainer, user) → transparency/accountability.

Engineering perspective

- Started discussions on FAT-MIR
Music recommender systems

- Based on the concept of similarity
  - User similarity
  - Artist similarity
  - Music content similarity

- Approaches

Music recommender systems

- Based on the concept of similarity
  - User similarity
  - Artist similarity
  - Music content similarity

- Approaches
  - Collaborative filtering: similar listeners

Music recommender systems

- Based on the concept of similarity:
  - User similarity
  - Artist similarity
  - Music content similarity

- Approaches:
  - Collaborative filtering
  - Music content description

Music recommender systems

- Based on the concept of similarity
  - User similarity
  - Artist similarity
  - Music content similarity

- Approaches
  - Collaborative filtering
  - Music content description
  - Music context description (web, lyrics, editorial metadata)

Music recommender systems

- Based on the concept of similarity
  - User similarity
  - Artist similarity
  - Music content similarity

- Approaches
  - Collaborative filtering
  - Music content description
  - Music context description
  - Hybrid

- Similarity vs diversity dilemma

Designing music recommenders

THE ABSTRACTION TRAPS IN DESIGNING SOCIOTECHNICAL SYSTEMS

1. **The Framing Trap**: Failure to model the entire system over which a social criterion will be enforced.

2. **The Portability Trap**: Failure to understand how repurposing algorithmic solutions designed for one social context may be misleading, inaccurate, or otherwise do harm when applied to a different context.

3. **The Formalism Trap**: Failure to account for the full meaning of social concepts which can be procedural, contextual, and contestable, and cannot be resolved through mathematical formalisms.

4. **The Ripple Effect Trap**: Failure to understand how the insertion of technology into an existing social system changes the behaviors and embedded values of the pre-existing system.

5. **The Solutionism Trap**: Failure to recognize the possibility that the best solution to a problem may not involve technology.


**Images from: [https://search.creativecommons.org/](https://search.creativecommons.org/)**
Changes in music listening

Music technologies are not neutral, they influence human perception and cognitive processes.

“We should be concerned about the loss of cultural diversity for the same reason that biologists worry about the loss of biodiversity: we don’t yet know what the loss will mean, but we do know that the loss will be irreversible.”


Filter Bubbles

Over-exposition to content which fits personal interests, hiding the diverse from the online experiences.

Echo Chambers

Tendency to relate mainly with like-minded people in online spaces, reinforcing polarization.

Cyberbalkanization

Appearance of online communities where frontiers shift from being geographical to being interests-based.

Develop a Framework for Defining and Evaluating Music Recommendation Diversity

Assess Music Recommendation Diversity

Understand the Consequences of Music Recommendation Diversity

Propose Countermeasures for Tuning Music Recommendation Diversity

https://lorenzoporcaro.wordpress.com/
Diversity of music recommender systems

The Specialties of Music Recommendation

very low consumption time in the dimension of minutes, whereas a book or a travel are consumed during days or weeks;

collection in sequences (e.g., playlists);

music often consumed passively (e.g., while jogging, travelling, working);

consumption is highly driven by situational context;

users are likely to appreciate the re-recommendation of the same item while a user is less likely to read the same news article over and over again;

music evokes strong emotions.

Future Directions and Visions in Music Recommender Systems Research

Psychologically-inspired music recommendation

Situation-aware music recommendation

Culture-aware music recommendation


Preliminary outcomes


Exploration of standard diversity measures from the Information Theory literature for performing comparative analysis of playlist datasets.

- Quantitative Approach
- Comparative Analysis (Historical/Technological)
- Playlist as a Static Object
- Information Theory / Information Retrieval background

**https://github.com/MTG/playlists-stat-analysis**
AOTM¹
# tracks: 972K
# playlist: 100K
type: user-generated
catalogue: user

CORN²
# tracks: 15K
# playlist: 75
type: radio playlist
catalogue: radio

SPOT³
# tracks: 2M
# playlist: 175
type: user-generated
catalogue: streaming

DEEZ⁴
# tracks: 227K
# playlist: 82K
type: user-generated
catalogue: streaming

⁴ Crawled in-house
Dataset Characterization

#1 - **Popularity**

- a. Track popularity as track frequency in the dataset
- b. Playlist popularity as average track popularity

*Simpson and Shannon indexes* ➔ measure of evenness between tracks popularity

*Gini coefficient* ➔ balance between playlists popularity

#2 - **Semantic Diversity**

- a. Semantic information from tag-embeddings
- b. Semantic distance between tracks as weighted sum of tag-distance
- c. Playlist diversity as average of tracks' pairwise tag-distance

*Descriptive statistics* ➔ playlist diversity trends

*Gini coefficient* ➔ balance between playlists diversity
Preliminary conclusions

→ Proposed metrics reflects differences between playlist datasets

- Streaming user-generated playlist datasets present a shorter long tail
- Radio playlists more (tag) diverse than user-generated playlists

→ Datasets biased towards Western culture (i.e. need for more non-Western playlist datasets!)

→ Software for playlist dataset analysis publicly available https://github.com/MTG/playlists-stat-analysis
Preliminary outcomes (ii)


First attempt of proposing new measures for evaluating the variations of recommendation lists in different listening scenarios.

- Qualitative Approach
- Mathematical Modelling of Variations
- Playlist as a Dynamic Object
- Set Theory / Calculus background
# Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>Integrative synthesis</th>
<th>Subordination service</th>
<th>Agonistic antagonistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of interdisciplinarity</td>
<td>Problem-solving</td>
<td>Practice-oriented</td>
<td>Other</td>
</tr>
<tr>
<td>Methodological orientations</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Interdisciplinarity sheet

<table>
<thead>
<tr>
<th>Disciplines</th>
<th>● Engineering and Technology (applied sciences)</th>
<th>● Music (humanities)</th>
<th>● Law (humanities)</th>
<th>● Sociology (social sciences)</th>
<th>● Psychology (social sciences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of interdisciplinarity</td>
<td>Integrative synthesis</td>
<td>Subordination service</td>
<td>Agonistic antagonistic</td>
<td>FAT-MIR</td>
<td></td>
</tr>
<tr>
<td>Methodological orientations</td>
<td>Problem-solving</td>
<td>Practice-oriented</td>
<td>Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Assessing the impact of AI on human behaviour: interdisciplinary views

Emilia Gómez
@emiliagogu

Joint work/slides from Vicky Charisi, Marius Miron, Songül Tolan, Nando Martínez-Plumed, Enrique Fernández-Macías, Annarosa Pesole, Maria Iglesias (JRC); Carlos Castillo, Lorenzo Porcaro (UPF); José H. Orallo (UPV); Luis Merino, Fernando Caballero (UPO); Bob Sturm (KTH); Emilio Gómez-González (US)