

DE LA RECHERCHE À L'INDUSTRIE

cea

Alexei Grinbaum

CEA-Saclay/LARSIM  
(France)

[www.cea.fr](http://www.cea.fr)



**IA**

# LA VALEUR ÉTHIQUE DU HASARD

# HUMAN AGENTS AND MACHINE VALUES

- ✓ RESEARCHER,  
PROGRAMMER,  
DESIGNER
- ✓ TRAINER
- ✓ PRODUCER
- ✓ OPERATOR
- ✓ USER



“... Nature rewards  
Perilous leaps. The prudent atom  
Simply insists upon its safety now,  
Security at all costs.”

W. H. Auden, *The Age of Anxiety*

HUMAN BENEFIT  
USEFULNESS  
LOYALTY  
RESPONSIBILITY

TRACEABILITY  
INTERPRETABILITY

SECURITY  
FAIRNESS  
TRANSPARENCY

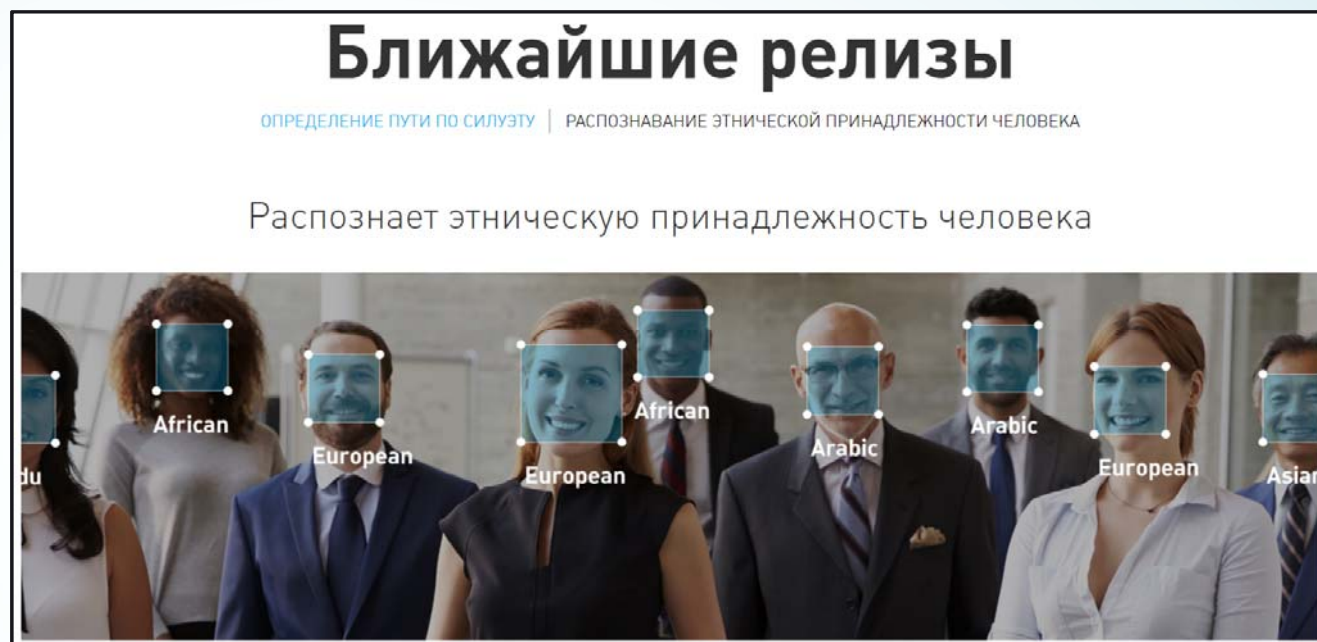
# “GOD’S MOVE”: NOT YET HUMAN



**“YESTERDAY, I WAS SURPRISED,” HE SAID.  
“BUT TODAY I AM SPEECHLESS.”**

**NEURAL NETWORK DETERMINES PROPERTIES, THEIR IMPORTANCE, AND THE CONNECTIONS BETWEEN THEM. IT “FORMULATES” ABOUT 80 PARAMETERS DESCRIBING FACIAL INFORMATION.**

**THE MEANING OF MANY OF THESE PARAMETERS IS UNKNOWN EVEN TO NTECH ITSELF.**



Next NTECH release: recognition of ethnic origin

“As a private person I value security more than privacy,” says Kukharenko.





## Facebook and YouTube should have learned from Microsoft's racist chatbot

- Facebook and YouTube have recently come under fire for offensive search suggestions.
- Microsoft made a Twitter chatbot in 2016 that was trained to say outrageous things by users, but Facebook and YouTube don't seem to have learned from the mistake.
- Psychological studies have shown that people are drawn to negative and offensive content, so engagement maximization drives the popularity of this content.

Cf. Staples example (S. Abiteboul)

### Factors leading to ethical problems

- Specification is difficult
- Training bias
- Learning without understanding
- Instability of learning
- Verification impossible, benchmarking difficult

CONFLICT  
OF VALUES





Cédric Villani

Être en mesure « d'ouvrir les boîtes noires » tient de l'enjeu démocratique.

If your argument is strong encryption no matter what, and we can and should create black boxes, that I think does not strike the kind of balance we have lived with for 200, 300 years...

Barack Obama



DE LA RECHERCHE À L'INDUSTRIE

cea

[www.cea.fr](http://www.cea.fr)

ETHICS, BY THEME  
ETHICS, VIA VALUES  
**ETHICS, THROUGH NARRATIVES**

9 mai 2018

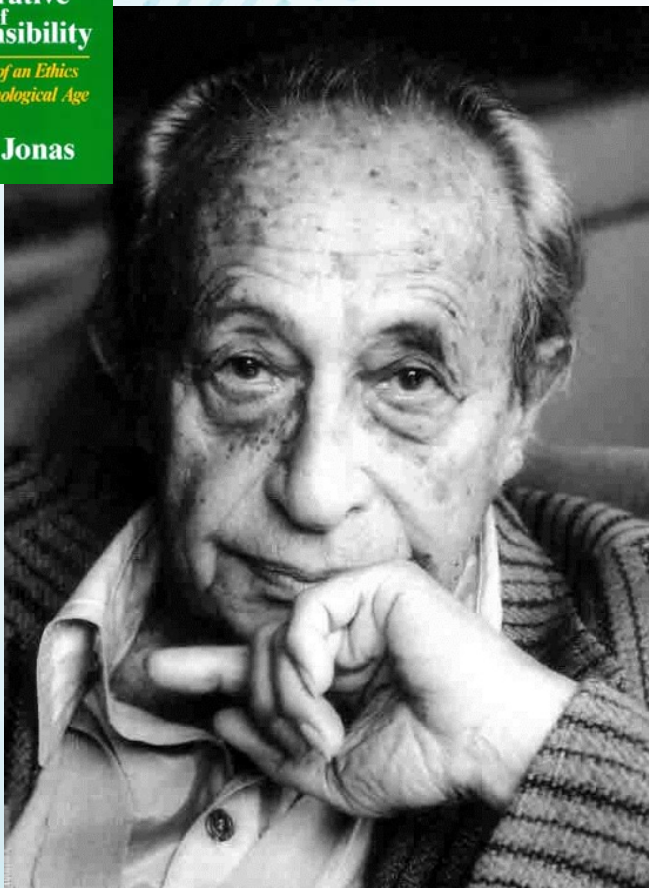


# HANS JONAS VS BARACK OBAMA

The  
Imperative  
of  
Responsibility

*In Search of an Ethics  
for the Technological Age*

Hans Jonas



“...the kind of balance we have lived with for 200, 300 years...”

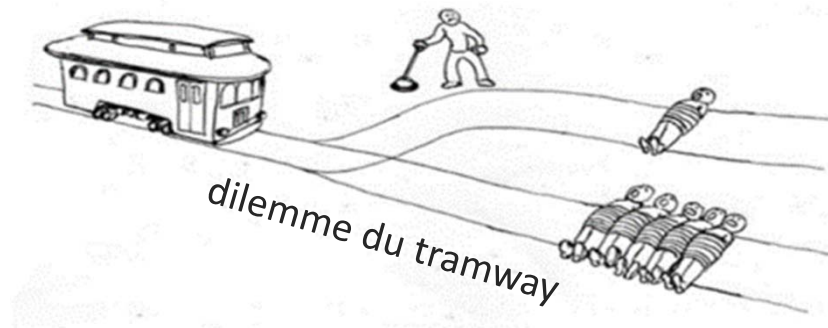
We, as members of technological societies, bear a historically unprecedented responsibility to future generations.

AG, "Satanas Ex Machina"  
(Desclée de Brouwer, 2018)



ÉTHIQUE HOMOLOGIQUE

## ÉTHIQUE ET POLITIQUE DANS LA CITÉ NUMÉRIQUE



**LA VALEUR DU HASARD.**

Les enfants d'Israël commirent une infidélité au sujet des choses vouées par interdit.

Josué déchire ses vêtements, se prosterne contre terre et reste devant l'arche de l'Alliance jusqu'au soir. (Js 7, 1-6)

Lorsque le Seigneur, béni soit-Il, dit à Josué: « Israël a péché », celui-ci Lui demanda : « Qui est celui qui a péché? »

**La réponse de l'Éternel : « Mais suis-je délateur ?  
Va et jette les dés. »**

Sur quoi, Josué alla et jeta les dés, et les dés tombèrent sur Achan. Celui-ci dit : « Josué, me condamnes-tu sans juste raison ? Toi et le prêtre Éléazar, vous êtes les deux hommes les plus grands de cette génération, or si j'avais à jeter les dés, ils auraient pu tomber sur l'un de vous. »

Josué répliqua : « Je te prie, ne médis pas des dés car Éretz Israël sera partagé par un tirage au sort. »

(Talmud de Babylone, Sanhedrin 43b)