**ESCOLA DE COMUNICAÇÕES E ARTES – USP**
Programa de Pós-Graduação em Ciências da Comunicação
Número e Algoritmo nas Interfaces Sociais da Comunicação

**Profª Drª Daniela Osvald Ramos**

# Algoritmos e Transparência

Edison Morais
Valdir Ribeiro Jr.

# 600TB/dia*

*Dados referentes a 2013.

Fonte: Blog do Facebook para programadores – http://bit.ly/2LgCXgD

# O que é ser transparente?

"Transparency is generally considered a **means to see the truth and motives behind people's actions** (Balkin 1999) and to ensure social **accountability and trust** (Breton 2006).

"On a very basic level, **transparency allows access to more information** which can influence **power relationships** between governments and citizens, business and customers, and in our case between news outlets and audiences (Bennis 2013)."

Diakopoulos e Koliska (2016)

# O que é ser transparente?

"Transparency reviews encourage the adoption of new or re-formatted informational production processes that produce information intended to fit the auspices of the review. In this way, **internal aspects of organizations are not 'made available'** but instead are re-oriented toward the production of specific forms of informational output that will externalize **(or make available) a particular version of the internal dynamics** of the organization"

Daniel Neyland (2007)

**TABLE 1**
Summary of transparency factors across four layers of algorithmic systems

| Layer | Factors |
|---|---|
| Data | • Information quality.<br>  ○ Accuracy.<br>  ○ Uncertainty (e.g. error margins).<br>  ○ Timeliness.<br>  ○ Completeness.<br>• Sampling method.<br>• Definitions of variables.<br>• Provenance (e.g. sources, public or private).<br>• Volume of training data used in machine learning.<br>• Assumptions of data collection.<br>• Inclusion of personally identifiable information. |
| Model | • Input variables and features.<br>• Target variable(s) for optimization.<br>• Feature weightings.<br>• Name or type of model.<br>• Software modeling tools used.<br>• Source code or pseudo-code.<br>• Ongoing human influence and updates.<br>• Explicitly embedded rules (e.g. thresholds). |
| Inference | • Existence and types of inferences made.<br>• Benchmarks for accuracy.<br>• Error analysis (including e.g. remediation standards).<br>• Confidence values or other uncertainty information. |
| Interface | • Algorithmic presence signal.<br>• On/off.<br>• Tweakability of inputs, weights. |

# Quando podemos ser transparentes?

Quando a abertura dos dados promove bem estar social. Exemplo:

– Dados agregados de segurança pública (NEV)

– Projeto serenata de amor

# Quando **não** devemos ser transparentes?

Quando há propriedade intelectual ou de Estado envolvida, ou ao tentar burlar o sistema, os indivíduos com acesso ao código seriam beneficiados. Podendo manipular o jogo em benefício próprio.

Exemplos:

– Algoritmo do Google

– Malha fina da Receita Federal

# Transparência

## Transparency

– Código fonte

## Accountability

– Responsabilidade, ética, prestação de contas

*"Accountability must be part of the system's design from the start. Designers of such systems - and the nontechnical stakeholders who often oversee or control system design must begin with oversight and accountability in mind."*

Kroll et al. (2016)

# Machine Learning

A Inspeção do código-fonte não é o suficiente para a compreensão da tomada de decisão automatizada.

O código, neste caso explicita apenas o método de machine learning utilizado e não a metodologia de tomada de decisão usada.

# Aleatoriedade e o Viés Cognitivo

Um algoritmo puramente randômico é um dos maiores desafios da computação.

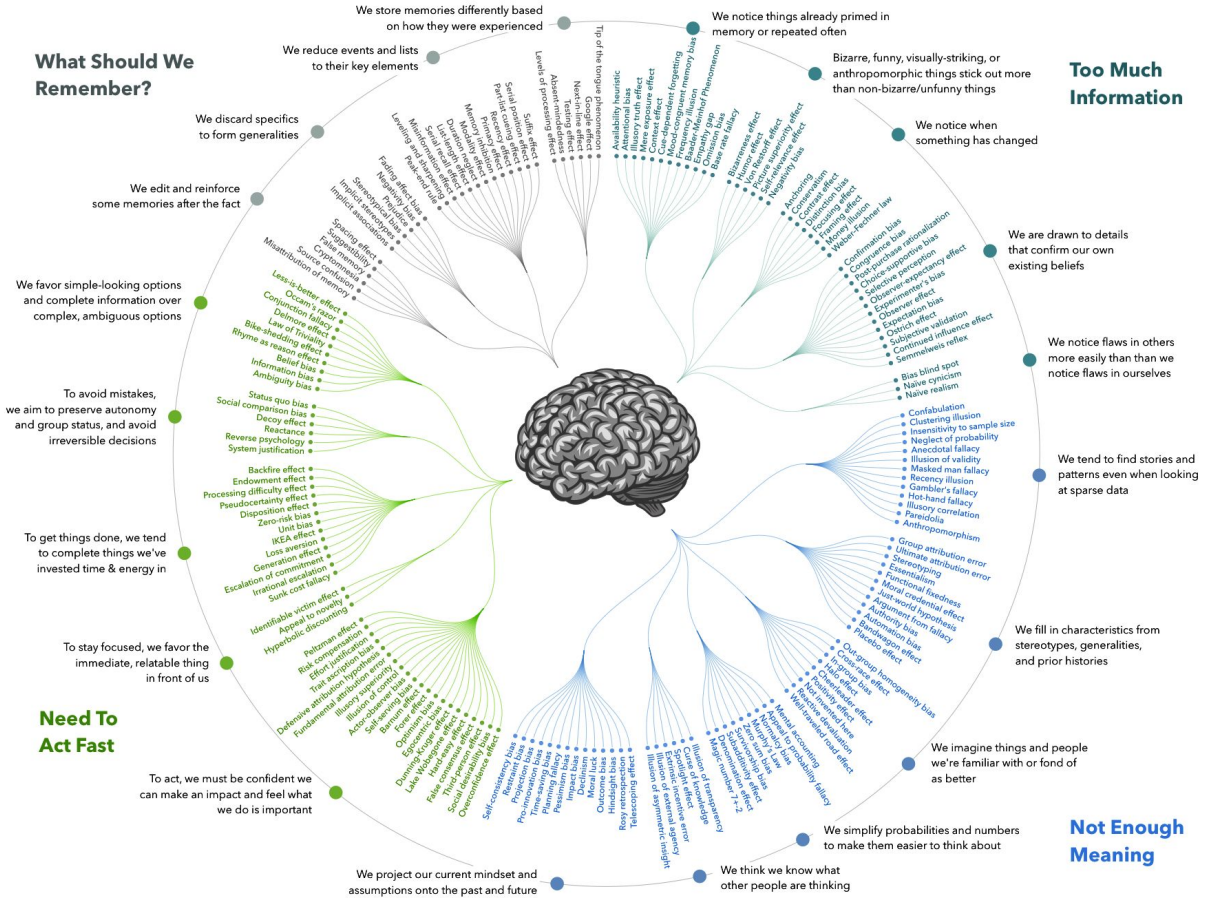
Processos randomizados garantem isenção e equidade de análise, removem vieses humanos

"Automation introduces a surprising benefit. By limiting the role of human discretion and intuition and relying upon computer-driven decisions this process protects minorities and others weaker groups."

Zarsky, 2012

Automated Predictions: Perception, Law and Policy

# COGNITIVE BIAS CODEX



# Perspectivas

É possível um algoritmo ser transparente no sentido "accountability" retornando a série de pré-requisitos que está respondendo.

Para isso é necessário ser projetado **desde o começo**, por desenvolvedores éticos.

Vellido, A, Martín-Guerrero, J. and Lisboa P. 2012. Making Machine leaning models intertepretable

# Referências bibliográficas


- Diakopoulos, N; Koliska, M. Algorithmic Transparency in the News Media (2016). Digital Journalism.

- Neyland, D. Achieving Transparency: The Visible, Invisible and Divisible in Academic Accountability Networks (2007)

- Joshua A. Kroll , Joanna Huey , Solon Barocas , Edward W. Felten , Joel R. Reidenberg , David G. Robinson & Harlan Yu Accountable Algorithms, 165 U. Pa. L. Rev. 633 (2017).

- Vellido, A, Martín-Guerrero, J. and Lisboa P. (2012): Making Machine leaning models intertepretable.

- Zarsky, Tal. Automated Prediction: Perception, Law, and Policy (2012). Communications of the ACM (Volume 55, Number 9), Sept. 2012, p. 33-35.