

Los problemas epistémicos de los sistemas de recomendación

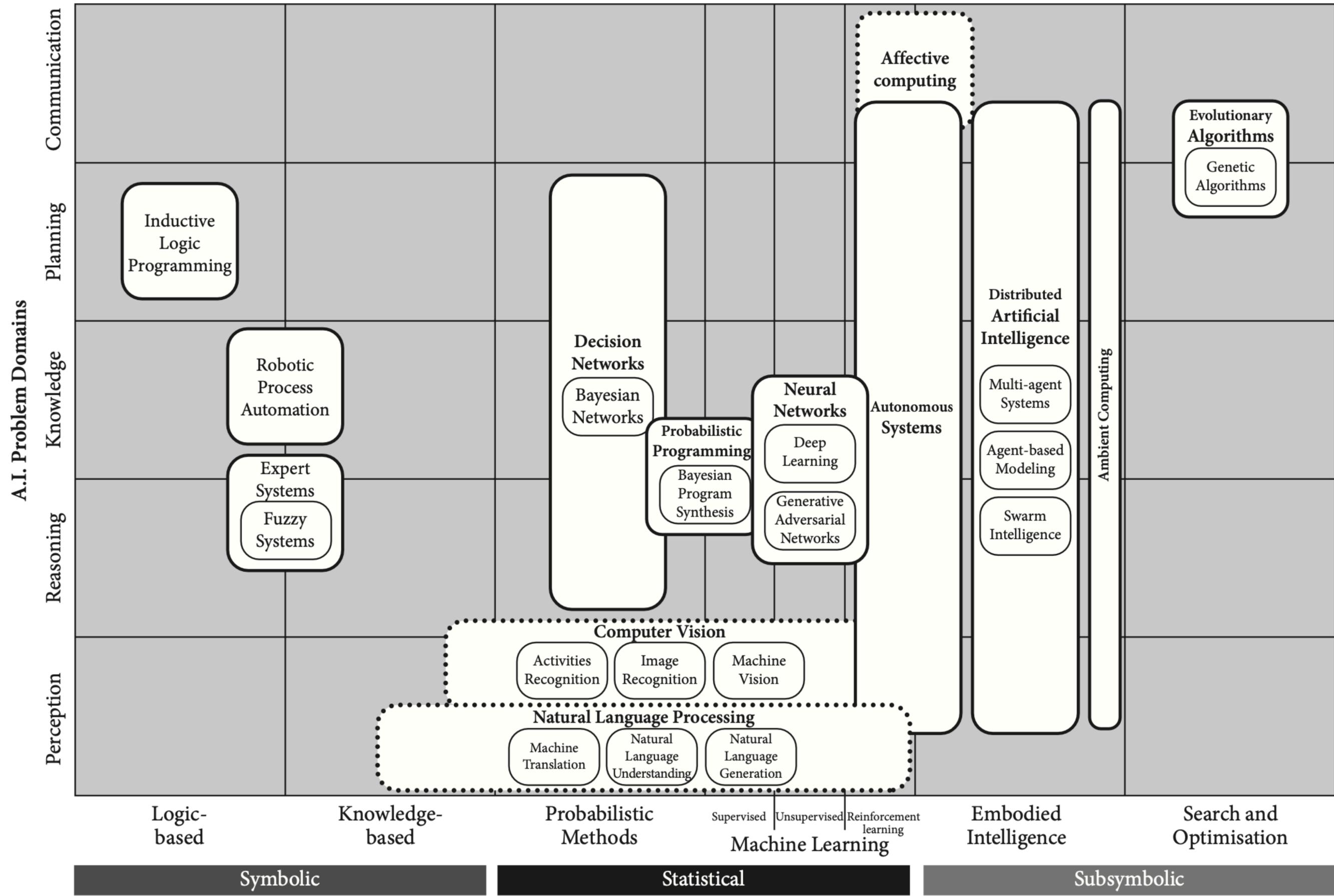
Entendiendo los desafíos de la interacción entre usuarios y algoritmos de personalización

En este trabajo analizo dos problemas epistémicos relacionados con la interacción entre usuarios humanos y sistemas de recomendación. Al primer problema lo llamo el problema de la pérdida de agencia -una situación donde un agente es silenciado, demeritado o reemplazado como sujeto epistémico (e.g., Floridi, 2023; Origgi & Ciranna, 2017). El segundo problema es lo que Wommer llama “insensibilidad agencial” (Woomer, 2017) -la falta de capacidades epistémicas para entender la evidencia. En conjunto, ambos problemas permiten mostrar que la mayor preocupación relacionada con los sistemas de recomendación no es que los usuarios se vuelven menos inteligentes al estar expuestos a la inteligencia artificial. En cambio, la preocupación es que la personalización algorítmica nos motiva a ser agentes epistémicos irresponsables, o en términos de Medina, a transformarnos en “ignorantes activos” (Medina, 2013, p. 39). Esta es una situación que debemos considerar si pretendemos resolver los desafíos que surgen a partir de nuestra interacción con la inteligencia artificial en plataformas digitales, al menos desde la perspectiva de los usuarios.

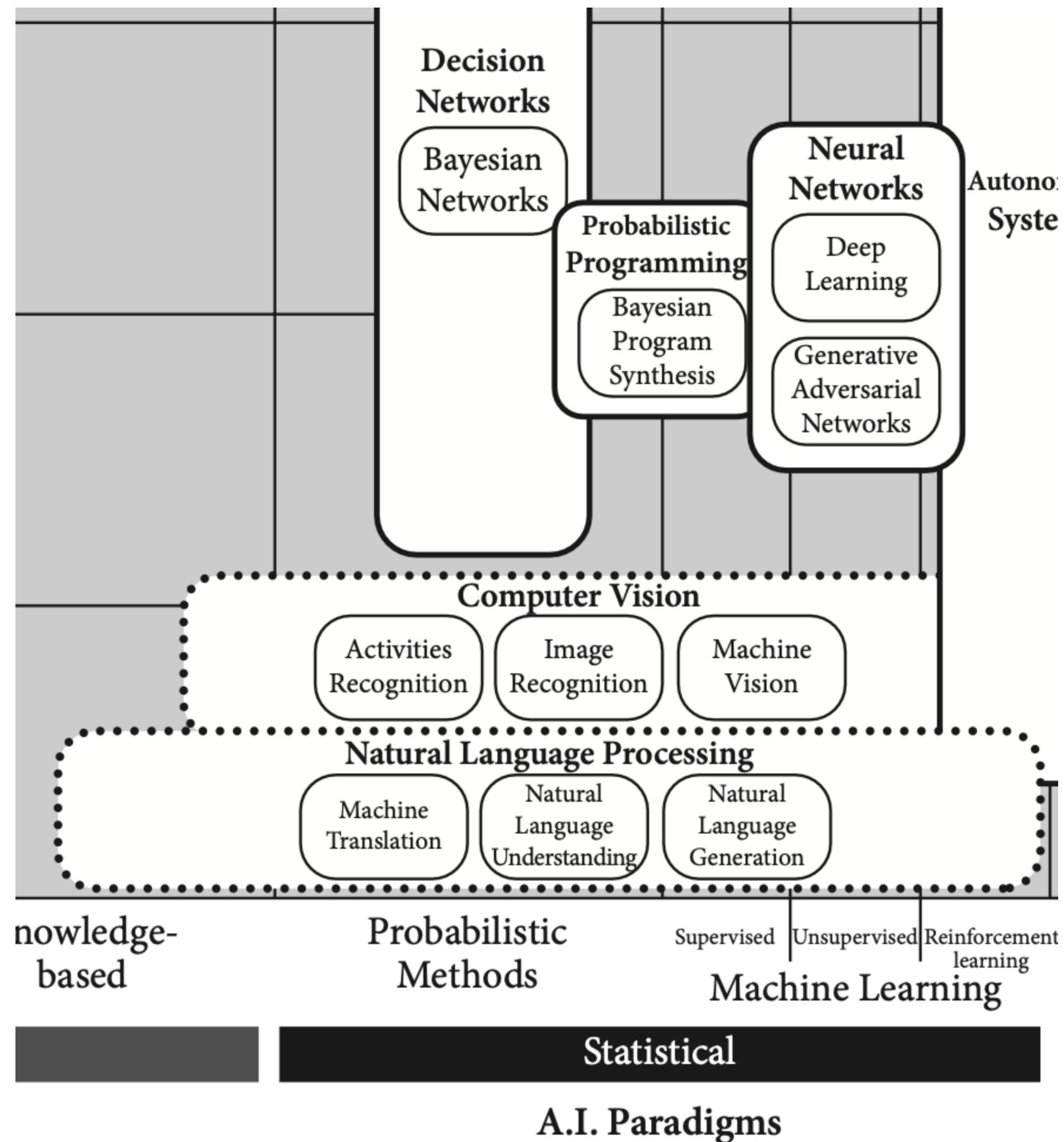


Our algorithm has detected that you're one of those annoying people who wants to know how our algorithm works. Our algorithm recommends that you be killed.





A.I. Paradigms



Corea, Francesco. Aug 29 2018. 'AI knowledge map: How to classify AI technologies, a sketch of a new AI technology landscape.' *Medium – Artificial Intelligence*, https://medium.com/@Francesco_Ai/Ai-Knowledge-Map-How-To-Classify-Ai-Technologies-6c073b969020.

“We have seen that successful AI is not about producing human intelligence but replacing it. A dishwasher does not clean the dishes as I do but, at the end of the process, its clean plates are indistinguishable from mine. Indeed, they may even be cleaner (effectiveness) while using fewer resources (efficiency).”

(Floridi, 2023).

- **La IA reproduce agencia, no inteligencia.**
- **La IA busca la reproducción de la conducta inteligente**

“The smart world obliquely consists of pervasive collections of data from a vast range of sensorised devices, infrastructures and environments. It puts us on a path where data about everything could be collected for smartness to function and flourish. The success of the smart world is consequently dependent upon vast scales of data collection, particularly from sensorised devices. The smart world is therefore the ‘collected’ world.”

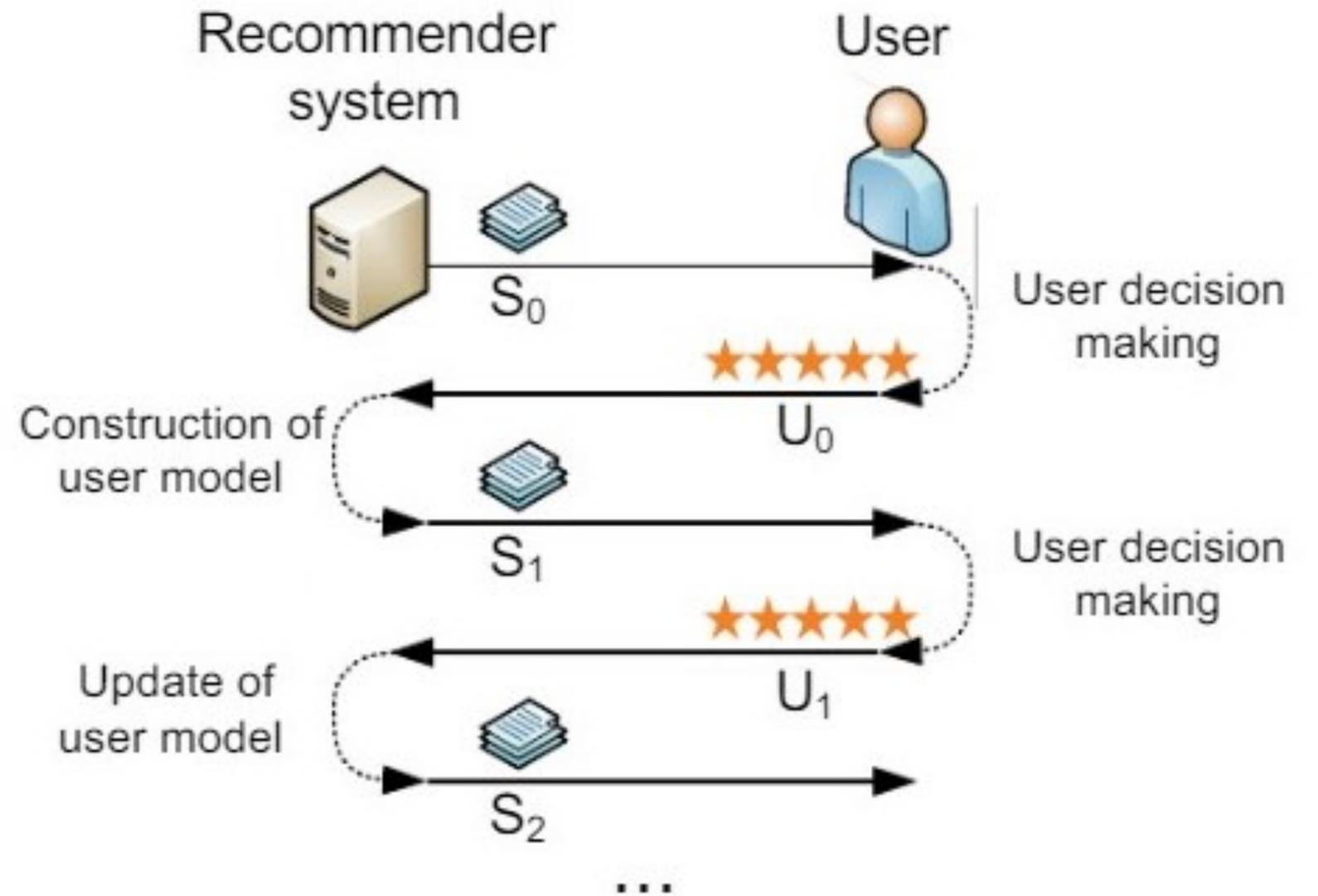
(Burdon, 2020).

- **Motores de búsqueda**
 - Resultados en función de las preferencias y experiencia pasada.
- **Sistemas de recomendación**
 - Noticias, música, personalización.
- **Sistemas de reputación**
 - Calificar usuarios, libros, películas, juegos.
- **Redes sociales**
 - Filtros en diferentes niveles

Sistemas de recomendación

¿Qué son?

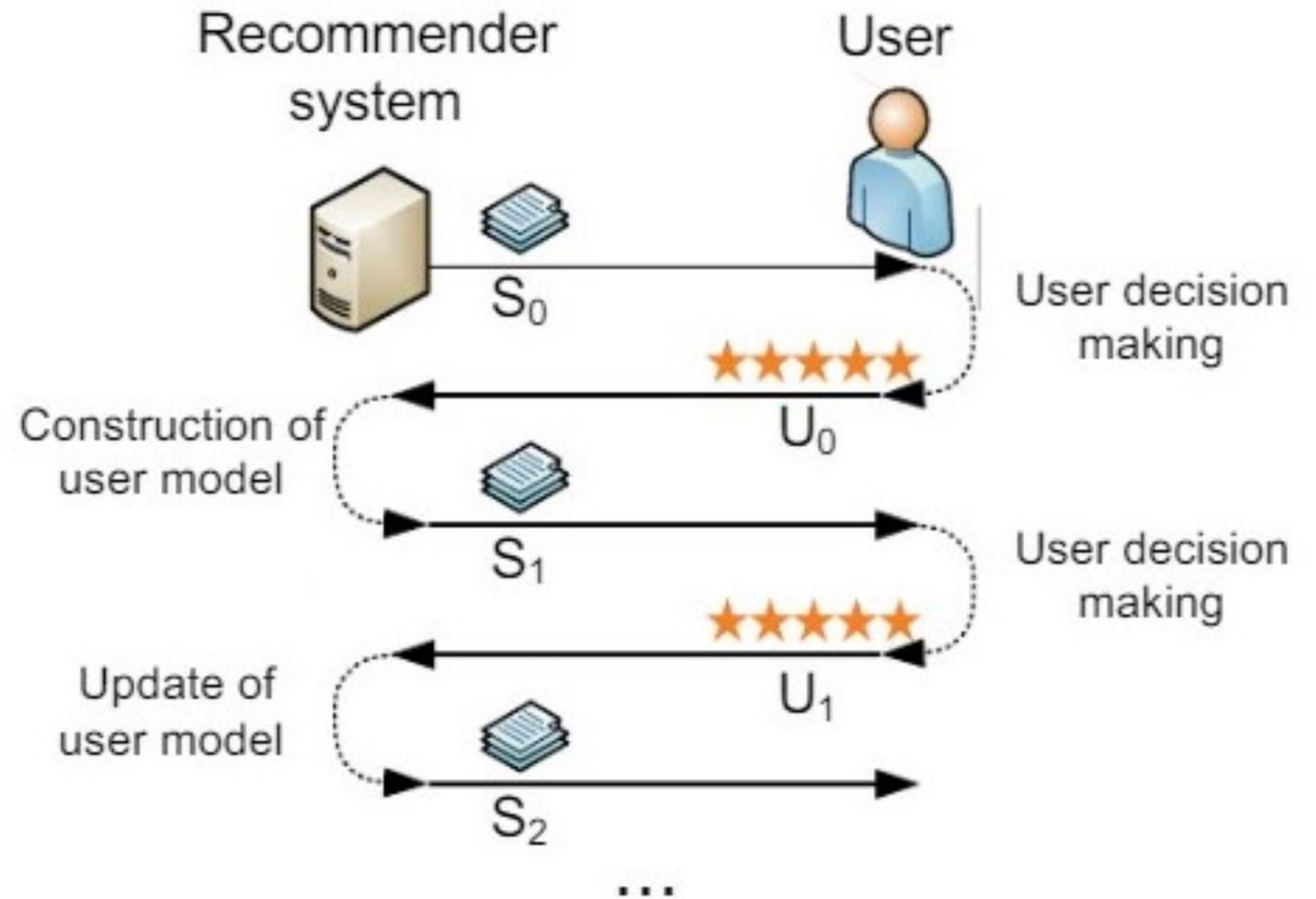
- Mecanismos de control y/o gestión de la información.
- Programas poseen un modelo de su entorno, el cual utilizan para adquirir más información y actualizar los parámetros de su diseño.



Sistemas de recomendación

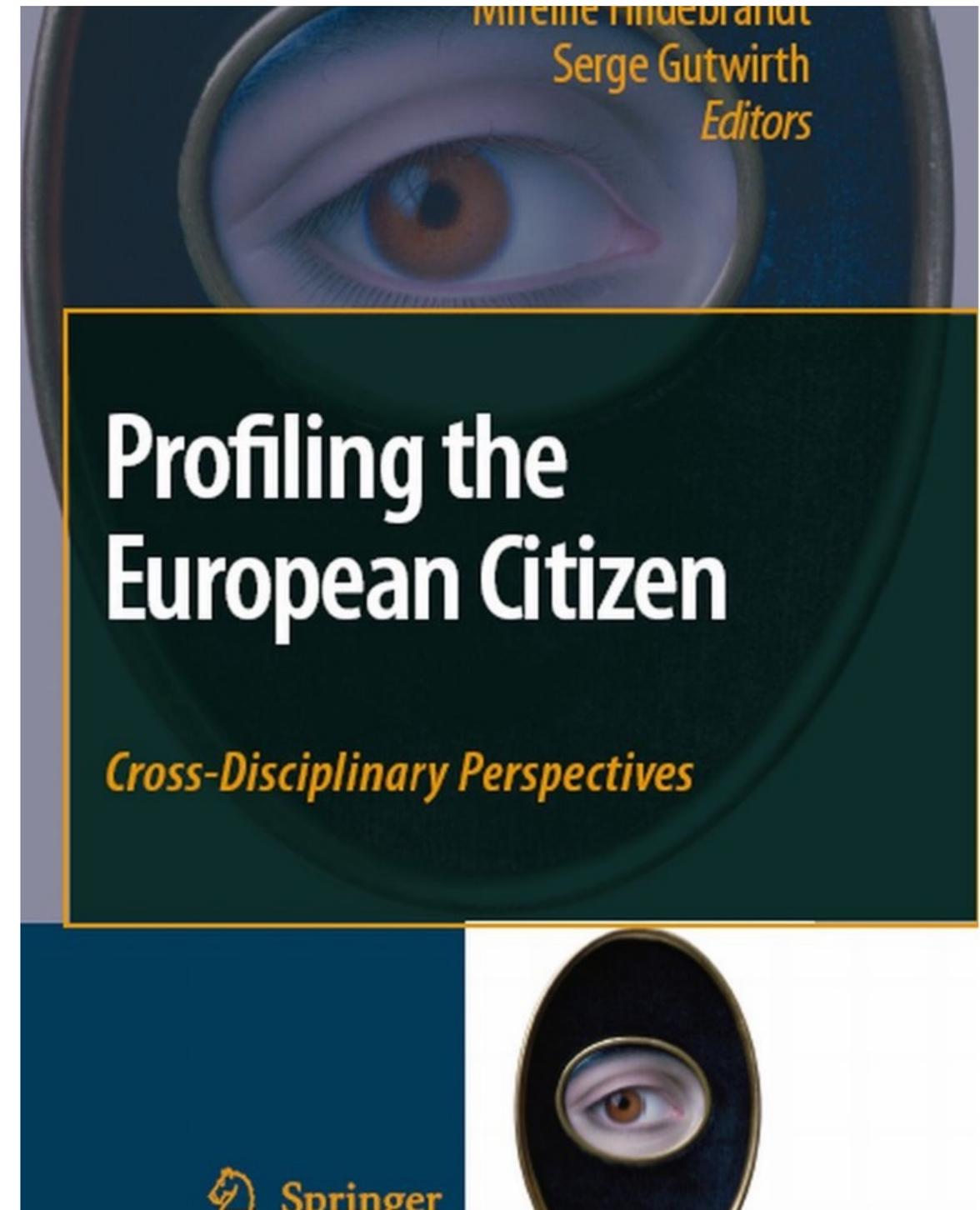
¿Qué son?

- “Cualquier programa que pueda describirse como teniendo un modelo de su entorno, el cual utiliza para realizar acciones que le permitan alcanzar sus objetivos, al mismo tiempo que adquiere más información que puede ser utilizada para actualizar los parámetros de su modelo” (Burr et al., 2018, p. 736).
- Lo que destaca es la interacción entre usuarios y agentes artificiales desde la lógica de la personalización.



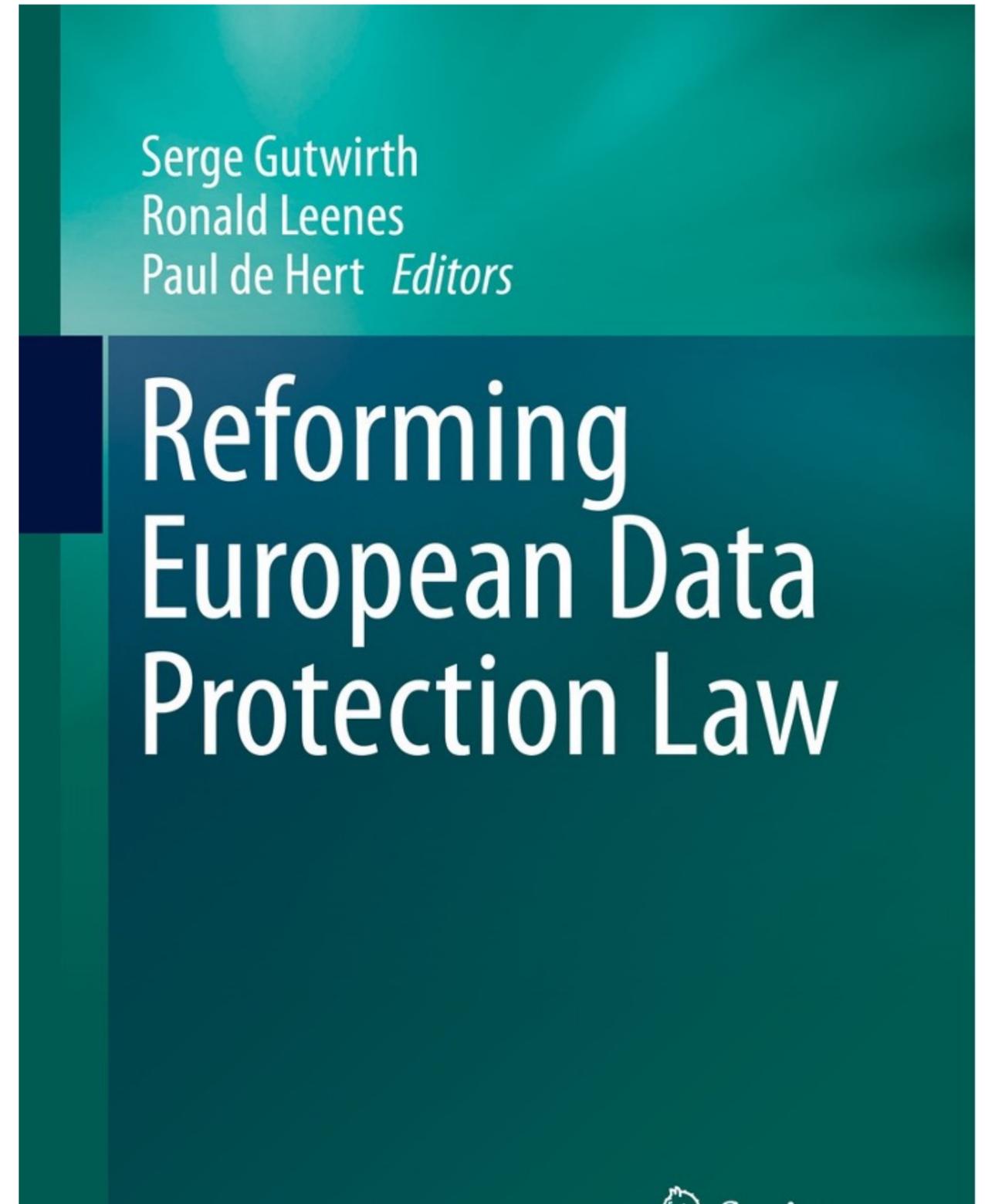
““El proceso de ‘descubrir’ correlaciones entre datos, en conjuntos de datos, que pueden ser utilizadas para identificar y representar un sujeto humano o no humano (individual o grupal) y/o la aplicación de perfiles (conjuntos de datos correlacionados) para individuar o representar un sujeto, o para identificar a un sujeto como integrante de un grupo o categoría.”

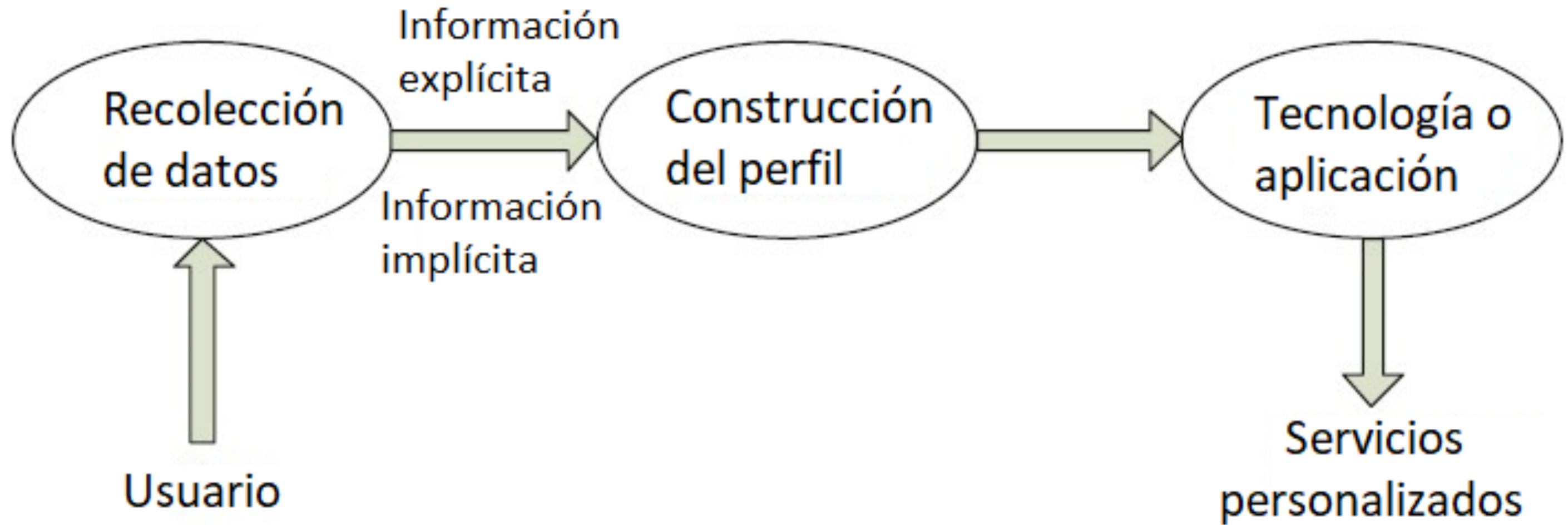
Hildebrand (2008, p. 19)



“El perfilamiento es una técnica de procesamiento automatizado de datos personales y no personales enfocada en producir conocimiento a partir de correlaciones inferidas desde datos en la forma de perfiles que pueden ser aplicados como la base para la toma de decisiones... Construir perfiles es el proceso de descubrir patrones desconocidos entre datos en grandes conjuntos de datos.”

Bosco et al. (2015, p. 8)



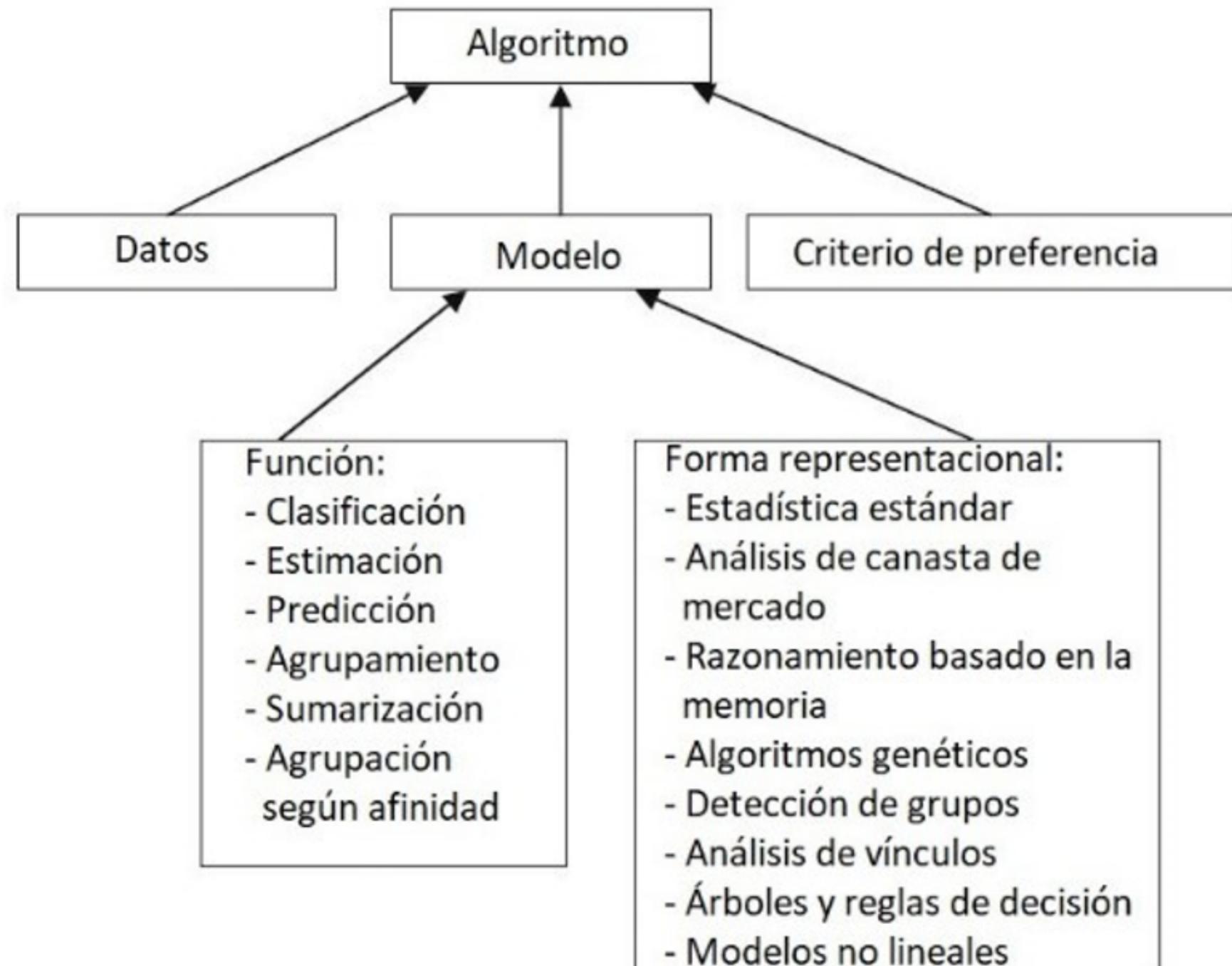


Construcción de perfiles de usuario para la personalización (Bozdog, 2013, p. 213)

Sistemas de recomendación

¿Qué son?

- Técnica de personalización basada en la minería de datos, el uso de algoritmos y la construcción de modelos que representan la conducta de los usuarios.
- Permiten el desarrollo de sistemas adaptativos (entornos digitales que aprenden la conducta de los usuarios).



Sistemas de recomendación

¿Los necesitamos?

- Nuestra experiencia digital implica enfrentar y resolver la sobrecarga de información
 - Nivel cognitivo: tenemos un límite en nuestras habilidades para procesar e interpretar la información.
 - Nivel técnico: existe un exceso de información de mala calidad que requiere el desarrollo de herramientas para evaluar la calidad de los contenidos.



Sistemas de recomendación

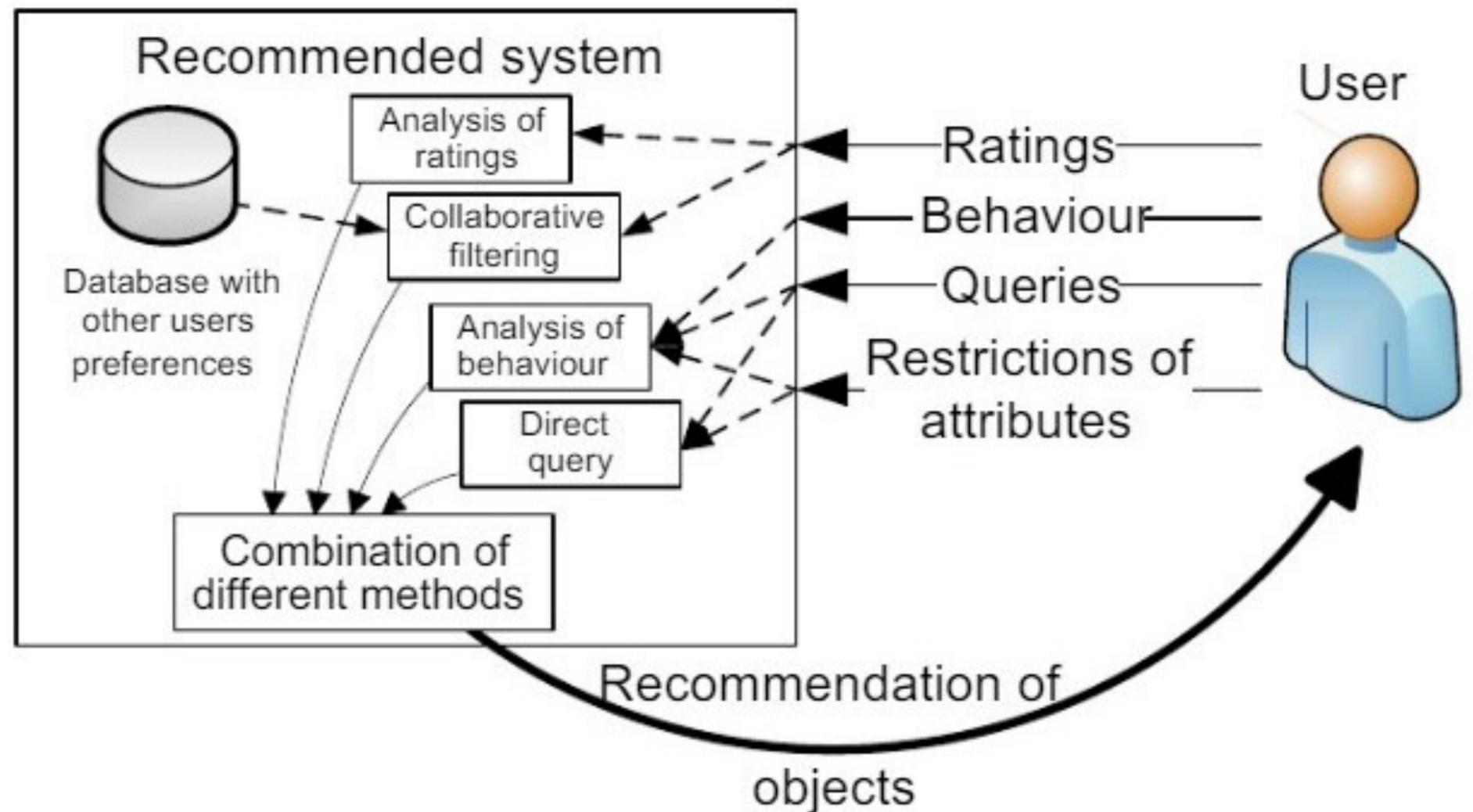
¿Los necesitamos?

- **Impacto directo:**

- Ofrecer incentivos.
- Mediante atajos heurísticos.
- Engañando o manipulando.
- Coercionando o restringiendo.

- **Impacto a largo plazo:**

- Cambio de creencias.
- Surgimiento de adicciones.
- Polarización



Estructura de un sistema de recomendación (Eckhardt, 2009, p. 61)

Tipos de personalización	¿Cómo operan?	Desafíos
Basada en redes sociales	Los usuarios entregan su información y ellos establecen los criterios básicos en función de sus preferencias. En adición los sitios utilizan la información almacenada para elaborar sus propios mecanismos de personalización	Los usuarios no tienen control respecto a cómo se utilizan sus datos. Los sitios suelen permitir que terceras partes accedan a los datos de los usuarios, lo cual supone una vulneración a la autonomía y la privacidad.
Perfilamiento conductual	Los sistemas digitales utilizan datos de diferentes fuentes para construir un modelo de la conducta de las personas de modo que los contenidos ofrecidos se ajustan a las predicciones del modelo. El proceso ocurre mayormente a espaldas de los usuarios.	La construcción de los modelos conductuales es opaca y ocurre a espaldas de los usuarios. No se requiere del consentimiento de una persona para utilizar sus datos
Basada en la ubicación	Descansa en el uso de dispositivos digitales con GPS u otro tipo de redes móviles. Permiten rastrear la ubicación de los usuarios y es mayormente utilizada con fines comerciales o de control	Existe la posibilidad que los programas revelen información sensible relativa a su ubicación o que terceras partes rastreen la actividad de los usuarios

Tipos de sistemas de recomendación	Descripción	Riesgos potenciales
Basados en el contenido	El sistema aprende a recomendar contenidos similares a los que una persona ha elegido en el pasado.	Los algoritmos utilizados pueden manipular la conducta de las personas ofreciendo incentivos para determinados tipos de contenidos
Filtrado colectivo	El sistema recomienda contenidos sobre la base de las preferencias de otras personas con gustos similares.	Homogeneización o despersonalización de los contenidos
Basados en la comunidad	El sistema recomienda contenidos sobre la base de las preferencias de otras personas con las que se posee algún tipo vínculo.	

El problema de la pérdida de agencia

Problema 1

- En términos de la experiencia digital, el diseño de entornos amigables para la IA implican un perjuicio no solo desde el conjunto de acciones posibles, sino también desde la perspectiva de la condición del usuario como sujeto de conocimiento.
- El diseño de los entornos digitales basados en la personalización excluye y/o margina a los usuarios como proveedores legítimos de información. Además, existe el riesgo de revelar información que un agente no está interesado en dar a conocer.

EPISTEMIC INJUSTICE

The case of digital environments

Gloria Origgi and Serena Ciranna

1. The Internet as a source of epistemic injustice

In her insightful and worldly acclaimed work on epistemic injustice, Miranda Fricker argues that people can be distinctively wronged in their capacity as knowers. Much of the discussion around the notion of *epistemic injustice* has revolved around power relations between different groups of people. In this chapter we would like to take a different perspective on epistemic injustice, by applying it to the context of human/ICT interactions. New technologies may be a source of epistemic harm by depriving people of their credibility about *themselves*. The massive gathering of big data about our own identity and behaviour creates a new asymmetry of power between algorithms and humans: algorithms are perceived today as *being better knowers of ourselves* than we are, thus weakening our entitlement to be credible about ourselves. We argue that these new cases of epistemic injustice are, under many aspects, *more centrally epistemic* than other cases described in the literature because they wrong us directly in our epistemic capacities and not only in our dignity as *knowledge givers*. The examples of epistemic harm we will discuss undermine our epistemic confidence about our self-knowledge, a kind of knowledge that has been considered for a long time as markedly different from all other kinds of knowledge because of its infallibility and self-presentness. We are diminished as knowers, especially in the most intimate part of our

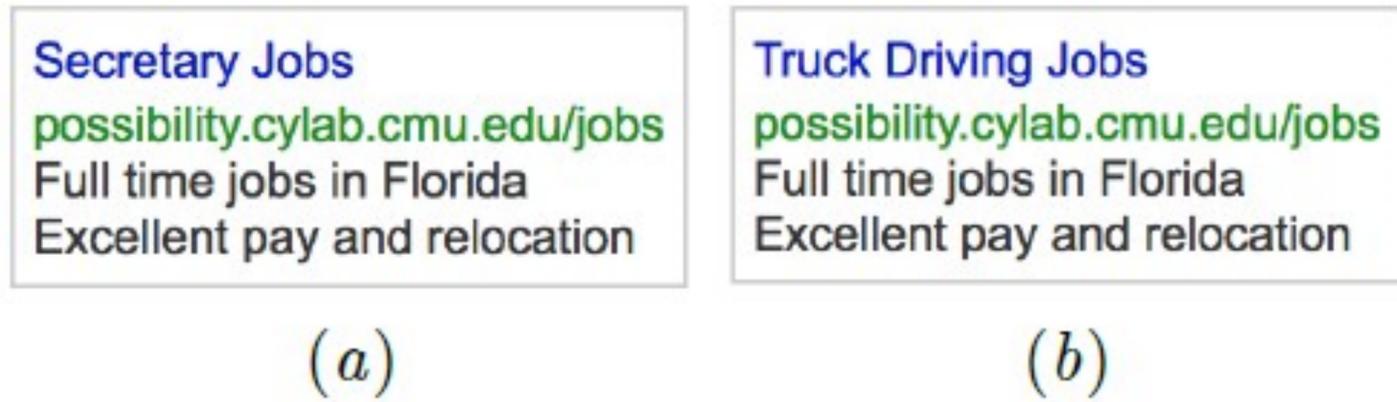


Figure 1: Ads approved by Google in 2015. The ad in the left (right) column was targeted to women (men).

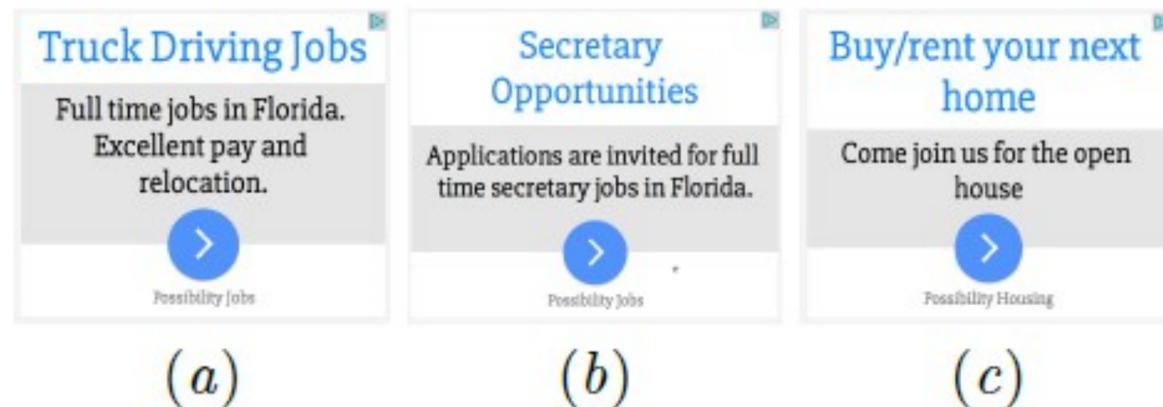
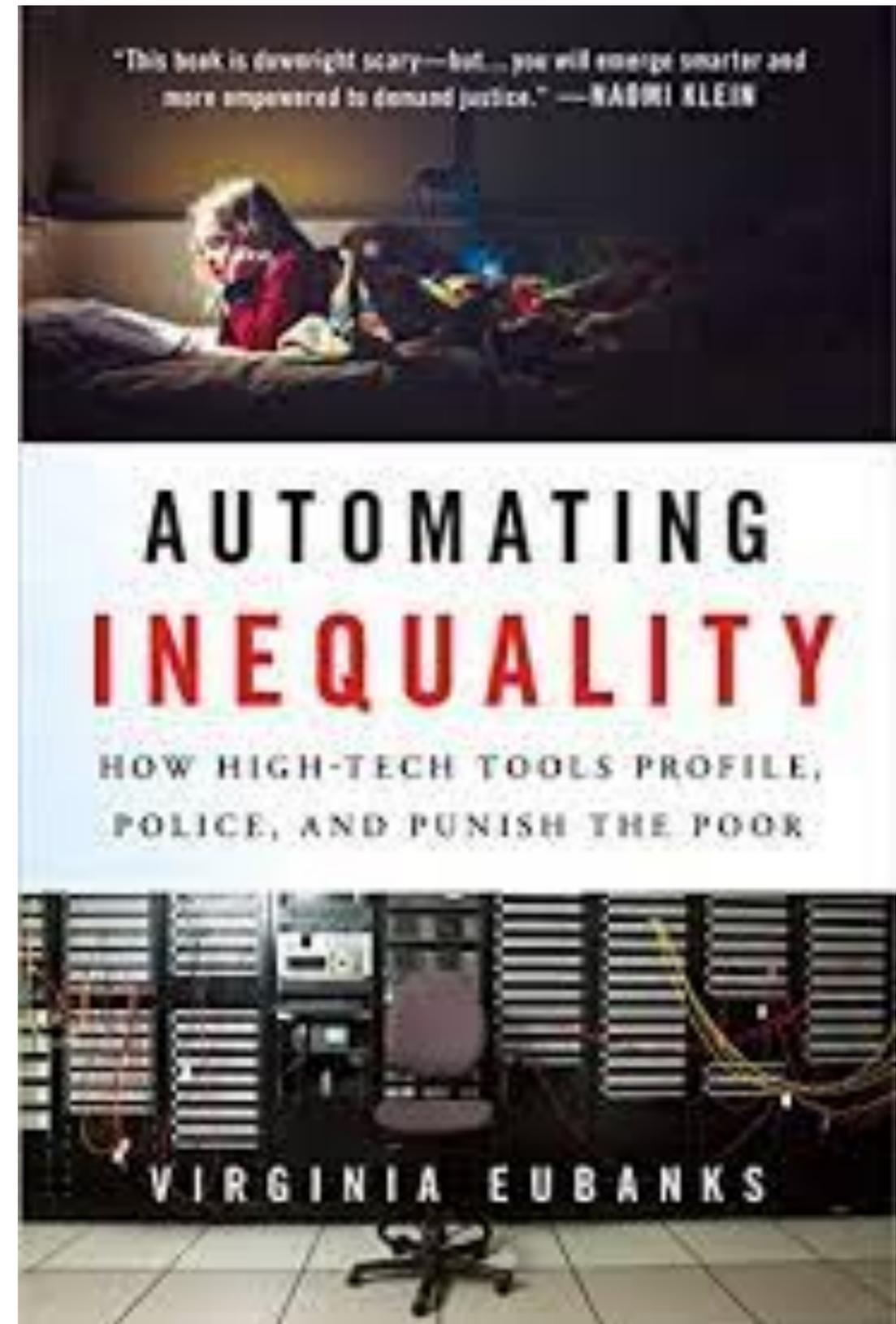


Figure 2: Ads approved and served by Google in 2017: truck driver jobs only to men, for secretary jobs only to women, for housing disparately.



La insensibilidad agencial

Problema 2

- Falla de un agente para utilizar adecuadamente las herramientas de acceso e interpretativas a su disposición.
 - Fallas de atención: ocurre cuando un agente no presta atención a la evidencia o los hechos disponibles y relevantes.
 - Falla de uso: Ocurre cuando un agente ha prestado atención a hechos o evidencia epistémicamente relevantes, pero falla en cambiar sus creencias en respuesta a la evidencia debido a una falla en el uso efectivo de sus herramientas interpretativas.



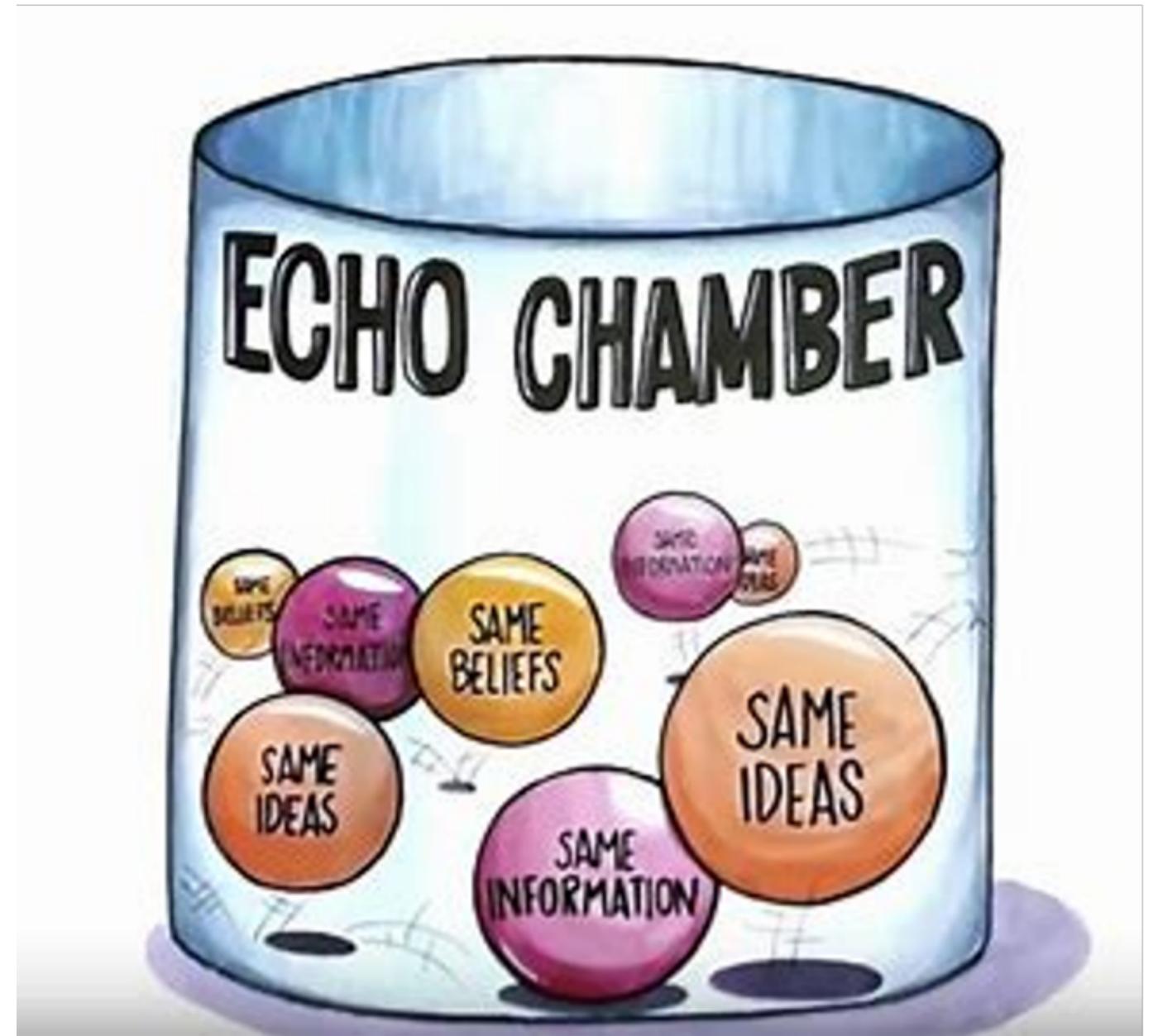
AGENTIAL INSENSITIVITY AND SOCIALLY SUPPORTED IGNORANCE

Published online by Cambridge University Press: 29 August 2017

Lauren Woomer



FILTER BUBBLES



CATALYST

feminism, theory, technoscience

ARTICLE

Amazon Echo and the Aesthetics of Whiteness

Thao Phan

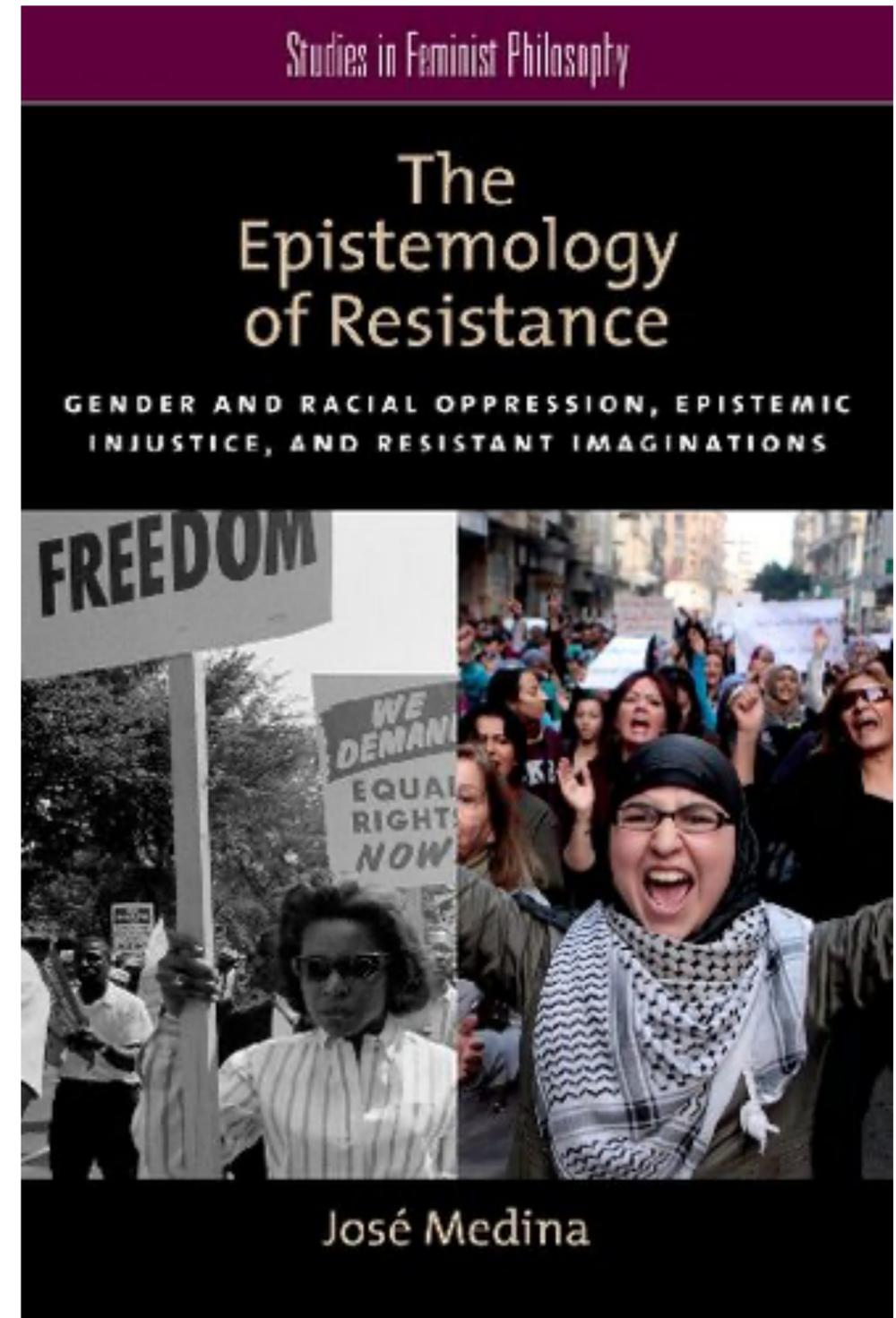
Alfred Deakin Institute for Citizenship and Globalisation, Deakin University

Thao.Phan1@deakin.edu.au



¿A dónde nos lleva esto?

“The epistemic vices of arrogance, laziness, and closed-mindedness greatly contribute to the production of a particular form of ignorance: active ignorance, an ignorance that occurs with the active participation of the subject and with a battery of defense mechanisms, an ignorance that is not easy to undo and correct, for this requires retraining—the reconfiguration of epistemic attitudes and habits—as well as social change. Those who are epistemically arrogant, lazy, and closed-minded are actively ignorant. Actively ignorant subjects are those who can be blamed not just for lacking particular pieces of knowledge, but also for having epistemic attitudes and habits that contribute to create and maintain bodies of ignorance.”



Cierre

- El desarrollo de las tecnologías de personalización y en especial el uso de perfiles de comportamiento, nos empuja a un estado de perjuicio informativo.
- Al mismo tiempo, suponen una transformación de los entornos digitales que incentiva la homogeneización de los usuarios y la asimetría informativa.
- La únicas soluciones parecen ser avanzar en la transparencia informativa y la distribución de la responsabilidad.
- ¿Qué podemos hacer como usuarios?



Referencias

- Bozdag, E., & van den Hoven, J. (2015). Breaking the Filter Bubble: Democracy and Design. *Ethics and Information Technology*, 17(4), 249-265. <https://doi.org/10.1007/s10676-015-9380-y>
- Floridi, L. (2023). *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford University Press. <https://doi.org/10.1093/oso/9780198883098.001.0001>
- Koene, A., Perez, E., Carter, C. J., Statache, R., Adolphs, S., O'Malley, C., Rodden, T., & McAuley, D. (2015). Ethics of Personalized Information Filtering. En T. Tiropanis, A. Vakali, L. Sartori, & P. Burnap (Eds.), *Internet Science* (pp. 123-132). Springer International Publishing. https://doi.org/10.1007/978-3-319-18609-2_10
- Medina, J. (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination*. Oxford University Press.
- Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender Systems and Their Ethical Challenges. *AI & SOCIETY*, 35(4), 957-967. <https://doi.org/10.1007/s00146-020-00950-y>
- Origgi, G., & Ciranna, S. (2017). Epistemic Injustice. The Case of Digital Environments. En I. J. Kidd, J. Medina, & P. Gaile Jr (Eds.), *The Routledge Handbook of Epistemic Injustice* (pp. 303-312). Routledge.
- Seaver, N. (2019). Captivating Algorithms: Recommender Systems as Traps. *Journal of Material Culture*, 24(4), 421-436. <https://doi.org/10.1177/1359183518820366>
- Woomer, L. (2017). Agential Insensitivity and Socially Supported Ignorance. *Episteme*, 16(1), 73-91. <https://doi.org/10.1017/epi.2017.28>