

# Introducción de la perspectiva de género en materias de inteligencia artificial

**Manuel Fernández Delgado**

Inteligencia Artificial Feminista

18 de noviembre de 2018

# Realidades (I)



- La inteligencia artificial (IA) puede mejorar la vida de las personas pero también puede ser perjudicial, sobre todo para las mujeres.
- Varias publicaciones alertan sobre la falta de neutralidad de los algoritmos de aprendizaje automático y sus prejuicios implícitos de género.
- Los sesgos de los sistemas informáticos son muy importantes porque la informática lo gestiona casi todo.

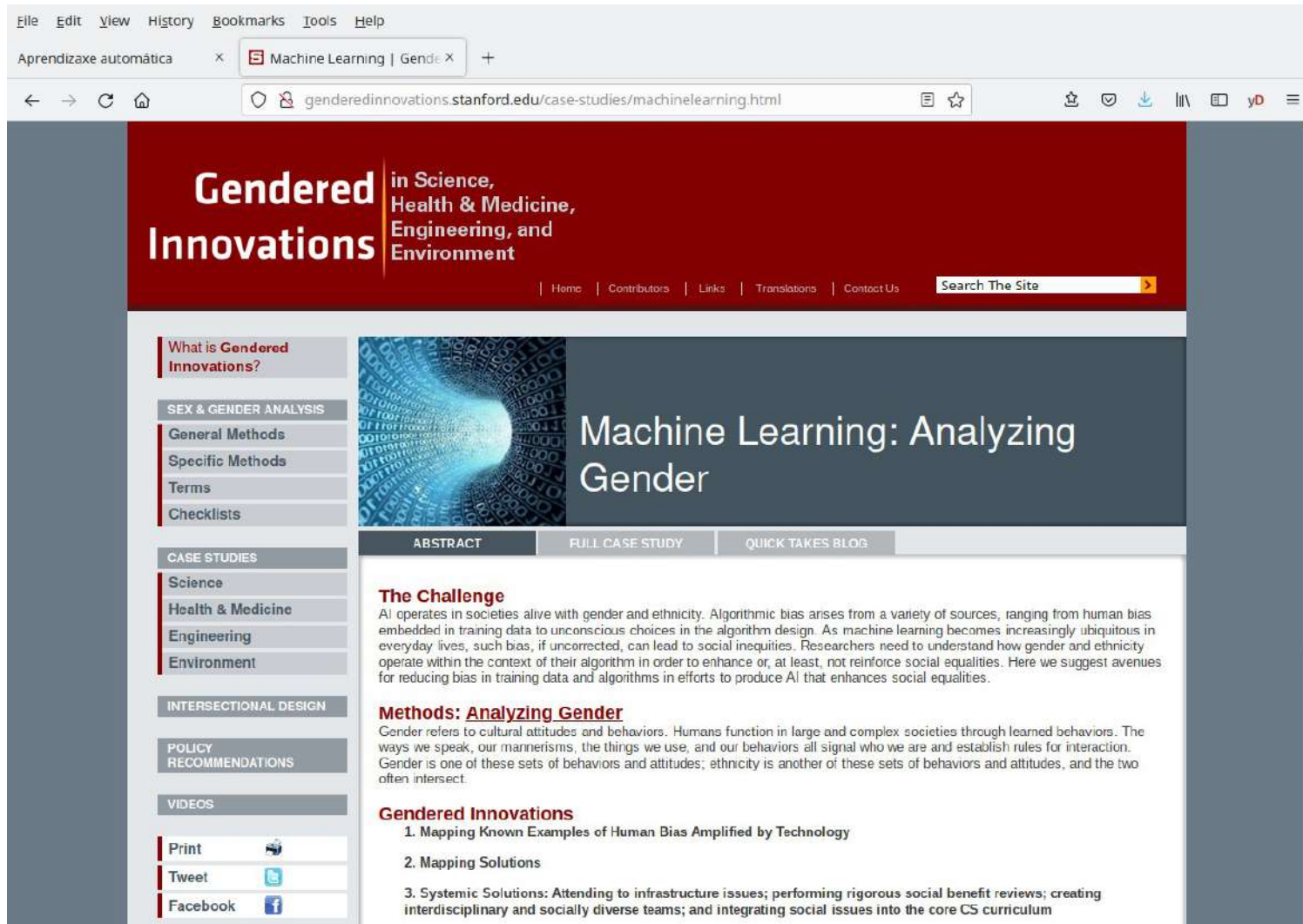
# Realidades (II)

- Aunque el 50% de la mujeres consumen tecnología, menos del 25% de mujeres trabajan en este campo.
- Es importante entender las fortalezas y limitaciones de la informática y la IA para no reproducir ni amplificar las desigualdades estructurales de la sociedad.
- Personalización informática “para lo que conviene”:
  - Sí, en buscadores de empleo con menores sueldos para mujeres.
  - No, para el lenguaje de género: “nombre de usuario”
- A menudo se culpa del sesgo de género a los algoritmos de IA, pero ...
- ... la IA es una parte muy pequeña de los sistemas informáticos.

# La IA como cortina de humo

- Muchos sistemas con sesgo de género no utilizan IA en absoluto.
- El sesgo de género se introduce en el diseño y la configuración del sistema informático.
- Atribuir el sesgo de género a los “algoritmos de IA” permite evadir la responsabilidad de ese sesgo.
- Esto no es nuevo: la informática siempre se usa para disculpar desatinos: “es por un error del programa”, “sería muy difícil tener en cuenta a todo el mundo”, etc.
- Esto proporciona una forma de evadir la culpabilidad.

# Ejemplo: proyecto de Gendered Innovations



The screenshot shows a web browser displaying the Gendered Innovations website. The browser's address bar shows the URL: `genderedinnovations.stanford.edu/case-studies/machinelearning.html`. The website header features the title "Gendered Innovations" in large white text on a dark red background, followed by "in Science, Health & Medicine, Engineering, and Environment" in smaller white text. Below the header is a navigation menu with links for "Home", "Contributors", "Links", "Translations", and "Contact Us", along with a search bar labeled "Search The Site".

The main content area is divided into a left sidebar and a main article. The sidebar contains several categories: "What is Gendered Innovations?", "SEX & GENDER ANALYSIS" (with sub-links for General Methods, Specific Methods, Terms, and Checklists), "CASE STUDIES" (with sub-links for Science, Health & Medicine, Engineering, and Environment), "INTERSECTIONAL DESIGN", "POLICY RECOMMENDATIONS", and "VIDEOS". At the bottom of the sidebar are social media sharing options for Print, Tweet, and Facebook.

The main article is titled "Machine Learning: Analyzing Gender" and features a blue-tinted image of binary code forming a tunnel. Below the title are three tabs: "ABSTRACT", "FULL CASE STUDY", and "QUICK TAKES BLOG". The "ABSTRACT" tab is selected, showing the following text:

**The Challenge**  
AI operates in societies alive with gender and ethnicity. Algorithmic bias arises from a variety of sources, ranging from human bias embedded in training data to unconscious choices in the algorithm design. As machine learning becomes increasingly ubiquitous in everyday lives, such bias, if uncorrected, can lead to social inequities. Researchers need to understand how gender and ethnicity operate within the context of their algorithm in order to enhance or, at least, not reinforce social equalities. Here we suggest avenues for reducing bias in training data and algorithms in efforts to produce AI that enhances social equalities.

**Methods: Analyzing Gender**  
Gender refers to cultural attitudes and behaviors. Humans function in large and complex societies through learned behaviors. The ways we speak, our mannerisms, the things we use, and our behaviors all signal who we are and establish rules for interaction. Gender is one of these sets of behaviors and attitudes; ethnicity is another of these sets of behaviors and attitudes, and the two often intersect.

**Gendered Innovations**

1. Mapping Known Examples of Human Bias Amplified by Technology
2. Mapping Solutions
3. Systemic Solutions: Attending to infrastructure issues; performing rigorous social benefit reviews; creating interdisciplinary and socially diverse teams; and integrating social issues into the core CS curriculum

# Ejemplo: interpretación de escenas

File Edit View History Bookmarks Tools Help

Aprendizaxe automática × google.com/url?sa=t&rct=j&q=× El automático traje del em: × +

← → ↻ 🏠 🔒 https://medium.com/bikolabs/el-automagico-traje-del-emperador-c2a0bbf6187b 📄 ☆ 📶 📧 📄 📄 yD ☰


simyo 13:32 81% simyo 13:32 81%

× ↶ ↷ 📌 Post × ↶ ↷ 📌 Post

**bikolabs**  
Un espacio para la creación, especulación y experimentación.

Follow

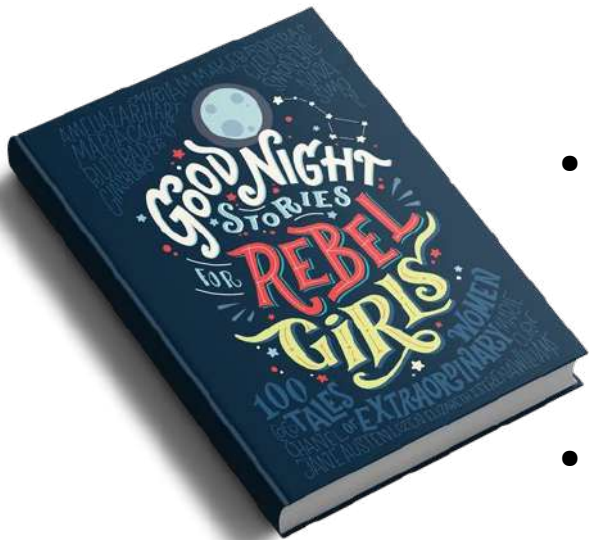
👏 10 💬 📌



**vestido de mujer azul y negro** 🔗

**polo de hombre a rayas azules y negras** 🔗

# Ejemplo: traductores automáticos



- Frase original es castellano: “Científicas, astronautas, levantadoras de pesas, juezas, chefs... cien ejemplos de determinación y audacia para las que sueñan en grande”
- Traducción al inglés con Google: “Scientists, astronauts, weightlifters, judges, chefs ... a hundred examples of determination and fearlessness for those who dream big”
- Vuelta a traducir al castellano con Google: “Científicos, astronautas, levantadores de pesas, jueces, cocineros ... un centenar de ejemplos de determinación y valentía para los que sueñan en grande”



# Perspectiva de género en la docencia universitaria

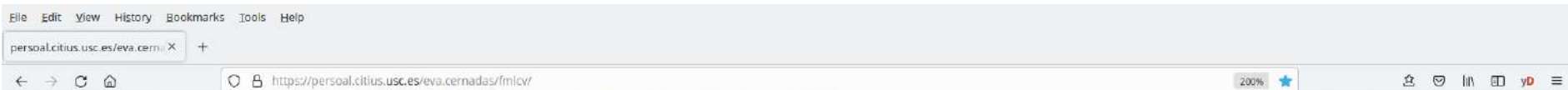
- La Unión Europea y España tienen un marco regulatorio para introducir la perspectiva de género en la educación superior.
- Actualmente, la introducción de la perspectiva de género es baja en las universidades españolas, e incluso inferior en los grados de Ingeniería Informática o Inteligencia Artificial.
- Se necesitan materiales y experiencias que inspiren al personal docente universitario.
- El equilibrio de género en la IA es fundamental para prevenir que los algoritmos perpetúen ideologías que discriminan a las mujeres.



# Nuestra experiencia en el curso 2020-2021

- **FMLCV: Fundamentals of Machine Learning for Computer Vision:**
  - Asignatura del International Master in Computer Vision, USC-UDC-UVigo-UPorto.
  - 14 estudiantes matriculados: 11 varones (78%) y 3 mujeres (22%).
  - Docencia telemática en inglés.
- **Aprendizaje automático:**
  - Materia optativa del Grado de Ingeniería Informática, USC.
  - 5 estudiantes matriculados: 1 mujer y 4 varones
  - Docencia presencial en galego.

# Página web materia FMLCV

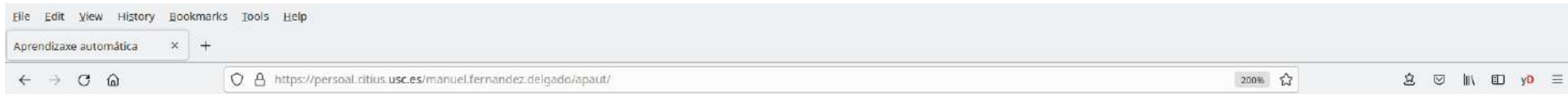


- a. R.O. Duda, P.E. Hart, D.G. Stork. **Pattern classification**. Wiley Interscience, 2000. ISBN: 978-0471056690
- b. Manuel Fernández-Delgado, Eva Cernadas, Senén Barro and Dinani Amorim, **Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?** Journal of Machine Learning Research, 2014.
- c. Other articles that will be recommended during the course.

## 2. Machine learning and gender:

- a. Criado Pérez, Caroline . **Invisible women**: Exposing data bias in a world designed for men. Random House, 2019 (english).
- b. Criado Pérez, Caroline . **La mujer invisible**: Descubre cómo los datos configuran un mundo hecho por y para los hombres. Seix Barral, 2020 (spanish).
- c. **El automático traje del emperador** (spanish). An experiment about the relation of the gender and the automatic algorithms in image understanding.
- d. **How to keep human out of AI** (english). Kriti Sharma explores how the lack of diversity in tech is creeping into our AI.
- e. **Machine intelligence makes humans morals more important** (english). Zeynep Tufekci explains how intelligent machines can fail in ways that don't fit human error patterns.
- f. **Gendered innovations** is a project promoted by Londa Schiebinger (you can see her biography in wikipedia) that harness the creative power of sex, gender, and intersectional analysis for innovation and discovery, and its importance in the research. It has sections for machine learning and robotics. For example the case of study **Face recognition**.

# Página web materia “Aprendizaje Automático”



## [Exercicios de clases interactivas](#)

### [Exame resolto 8 abril 2021](#)

#### **Aprendizaxe automática, xénero e ética:**

- [How to keep human out of AI](#). Vídeo de conferencia TED (en inglés) onde Kriti Sharma explica como a falta de diversidade dexenera a Intelixencia Artificial
- [El automágico traje del emperador](#). Artigo sobre a relación entre aprendizaxe automática e xénero no recoñecemento de imaxes.
- Artigo [“Científicos ilustran con un ejemplo simple el peligro de la inteligencia artificial que puede "destruir la humanidad”](#), Russia Today, 26 decembro 2020
- [Machine intelligence makes humans morals more important](#) (vídeo de conferencia TED, en inglés). Zeynep Tufekci explica que os algoritmos de aprendizaxe automática fallan de forma distinta ás persoas.
- Artigo [“Hacia la singularidad: Inteligencia Artificial dominando el mundo”](#). Blog de Juan Pérez Ventura, <http://vaventura.com>, 24 marzo 2018.
- [Gendered innovations](#) é un proxecto promocionado por [Londa Schiebinger](#) sobre a importancia do xénero e a análise intersectorial na innovación e investigación, especialmente en aprendizaxe automática, robótica e recoñecemento facial.
- Petición: [“¡Prohibir el reconocimiento facial en Europa!”](#). Resume dos inconvintes do recoñecemento facial.
- Artigo [“Petición de moratoria en el uso de sistemas de reconocimiento facial ante sus posibles riesgos”](#), Diario “El Español”, Autor: Alfonso Muñoz, 25 de marzo de 2021, 01:23h

# Actividades para introducir la perspectiva de género

- Aspectos a considerar cuando se planifica el curso:
  - Discusión de cuestiones de género y ética de los sistemas inteligentes en el aula (foro en FMLCV, participación en clases para ApAu).
  - Uso de lenguaje inclusivo en la oratoria y en los documentos escritos.
  - Visibilización del trabajo de las mujeres en IA.
  - Gestión del aula: favorecer y realzar opiniones de las alumnas.
  - Seleccionar ejemplos que promuevan la igualdad.
  - Cuestionar el sistema de evaluación.

# Visibilización: reseña de mujeres importantes en la Informática

File Edit View History Bookmarks Tools Help

operadores\_relacionais\_loxicos X +

← → ↻ 🏠 [https://persoal.citius.usc.es/manuel.fernandez.delgado/informatica/material/fortran/operadores\\_relacionais\\_loxicos.pdf](https://persoal.citius.usc.es/manuel.fernandez.delgado/informatica/material/fortran/operadores_relacionais_loxicos.pdf) ☆

🔍 ↻ 🖨️ 📄 🗑️ 🏠

1 of 8 Automatic Zoom

## Programadoras da NASA no programa espacial de viaxe á lúa (1969)

Mary Jackson: enxeñeira da NASA



Katherine Johnson: **matemática** e programadora da NASA  
Primeira civil en recibir a **Medalla de Ouro** do Congreso USA



Dorothy Vaughan: programadora en **Fortran** do primeiro ordenador da NASA  
Programación estruturada en Fortran

Película  
"Figuras ocultas"

Operadores relacionais e lóxicos 1



# Comentarios del alumnado

- Para eliminar los prejuicios de los algoritmos de aprendizaje automático tendremos que eliminarlos de nosotr@s mism@s.
- Los prejuicios de género deberían ser nuestra principal preocupación cuando desarrollamos sistemas de IA.
- La recopilación de datos de las redes para su uso en los algoritmos de aprendizaje automático debe ser monitorizado por un grupo diverso de personas para evitar reproducir estereotipos de género.
- Otros asuntos de tipo ético y moral como: ¿debería la legislación prohibir el uso del reconocimiento facial para sistemas de vigilancia?.

# Conclusiones de la experiencia

- Elevada participación y sensibilidad de las alumnas hacia las cuestiones de género y éticas en IA.
- Sorprendente nivel ético de una parte del alumnado, principalmente alumnas.
- Se destacó la importancia de considerar la diversidad en el diseño de sistemas de aprendizaje automático.
- Elevada participación (superior al 50%) y sensibilidad hacia los peligros de la IA.



# Reflexión personal

**“No seré cómplice de la discriminación hacia las mujeres en la inteligencia artificial, ni en las universidades ni a consecuencia de los sistemas informáticos que creamos”**