

Detectores de mentiras en las fronteras, ¿en serio?

Happy stats hour. Societat Catalana D'Estadística

Javier Sánchez-Monedero

jsanchezm at uco dot es

18 marzo 2021

Universidad de Córdoba y Data Justice Lab



datajusticelab.org

datajusticeproject.net

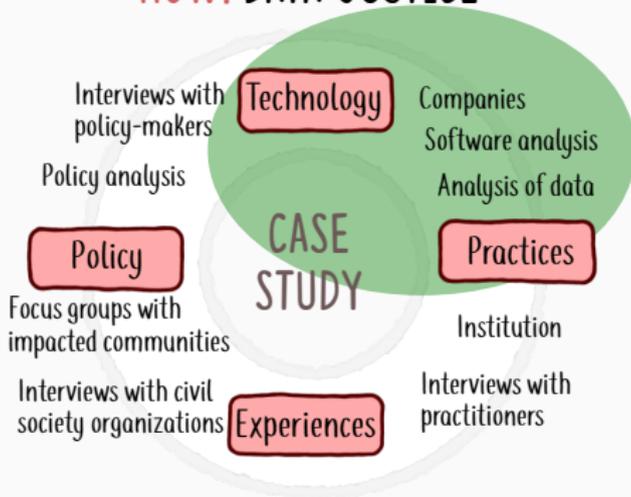
Data Justice

Analizando el impacto social de los datos

2016: Más allá de la privacidad

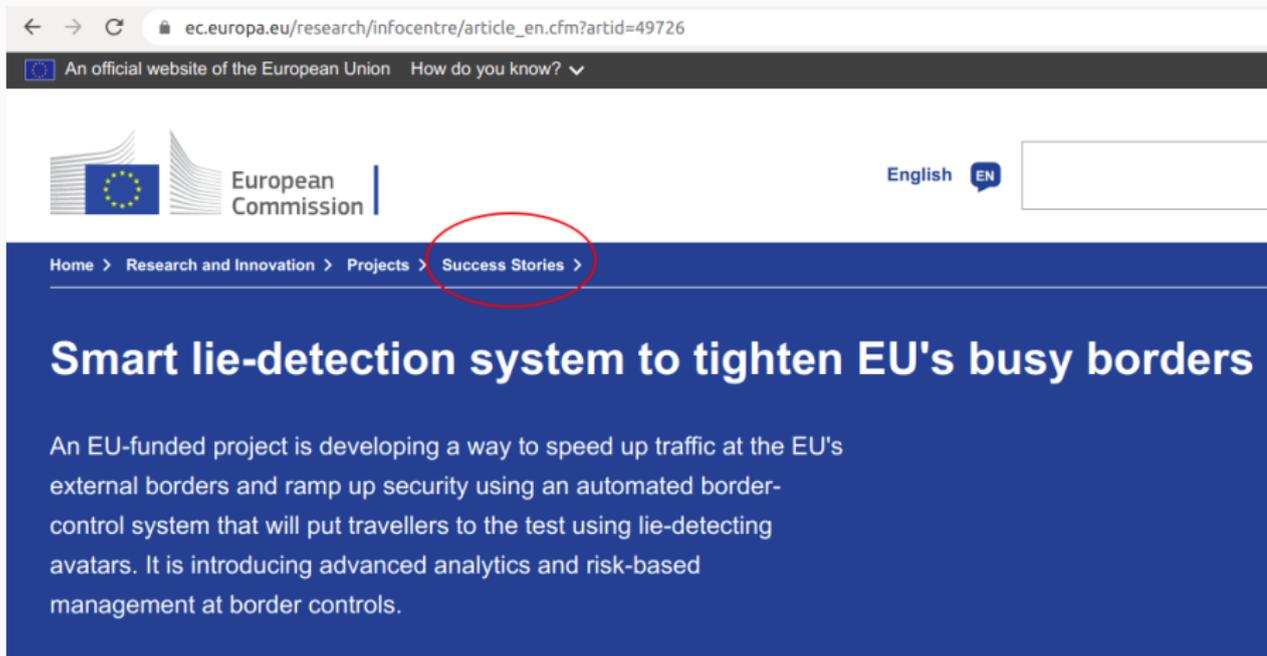
2020: Más allá de los sesgos algorítmicos

HOW? DATA JUSTICE



<https://datajusticeproject.net/about>

Fronteras engañosas



The image shows a browser window displaying an article on the European Commission website. The address bar shows the URL `ec.europa.eu/research/infocentre/article_en.cfm?artid=49726`. The page header includes the European Commission logo and the text "European Commission". A language selector shows "English" with a speech bubble icon. A breadcrumb trail at the top of the main content area reads "Home > Research and Innovation > Projects > **Success Stories** >". The main heading of the article is "Smart lie-detection system to tighten EU's busy borders". The introductory text describes an EU-funded project aimed at speeding up traffic and enhancing security at EU borders through an automated border-control system using lie-detecting avatars and advanced analytics.

ec.europa.eu/research/infocentre/article_en.cfm?artid=49726

An official website of the European Union How do you know? ▾

European Commission

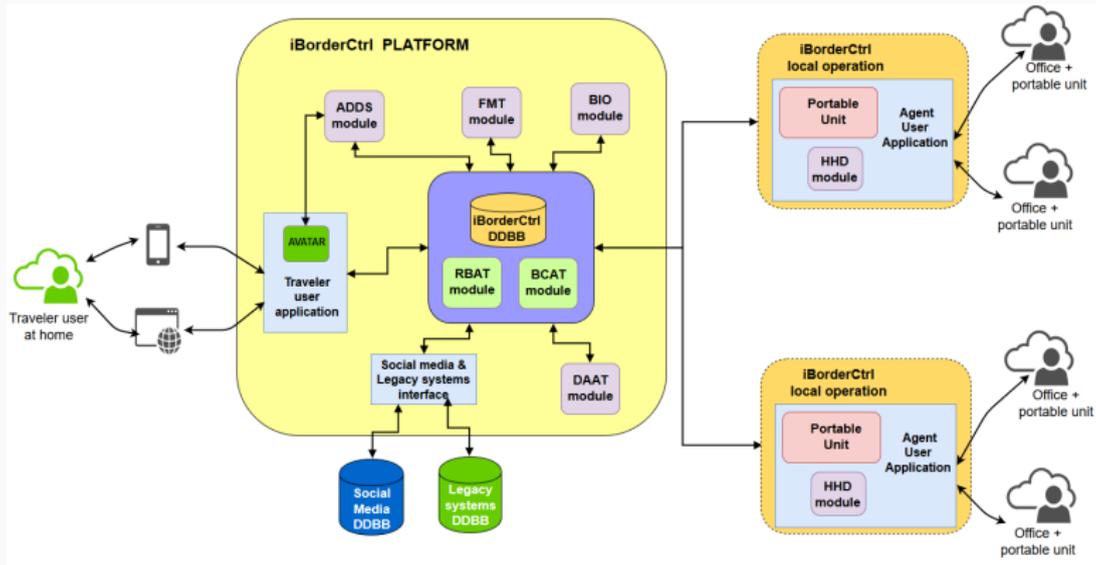
English EN

Home > Research and Innovation > Projects > **Success Stories** >

Smart lie-detection system to tighten EU's busy borders

An EU-funded project is developing a way to speed up traffic at the EU's external borders and ramp up security using an automated border-control system that will put travellers to the test using lie-detecting avatars. It is introducing advanced analytics and risk-based management at border controls.

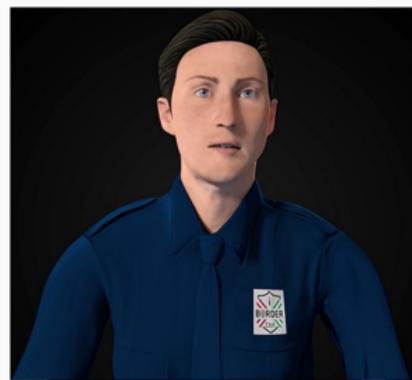
iBorderCtrl



iBorderCtrl (Intelligent Portable Control System, iborderctrl.eu)

iBorderCtrl (Intelligent Portable Control System):

- Control de viajeros **(no-UE) y migrantes**
- Financiado por el programa H2020 (4.5M€)
- Cruce de fronteras en **dos pasos**:
 - Registro previo al viaje desde casa
 - Entrevista automática con un agente virtual en la frontera
- Evaluación de **“riesgo”** automática
- **Detección de engaños/mentiras** a través de análisis facial (‘biomarkers of deceit’)
- En función del **riesgo y el engaño estimados**, la persona será entrevistada por un agente humano
- **Pilotos en Hungría, Grecia y Letonia** en 2018



¿Detectores de mentiras?



Los detectores de mentiras **no tienen validez científica** [National Research Council \[2003\]](#):

- Idea base: hay una serie de respuestas fisiológicas universales que una persona produce involuntariamente como resultado de mentir.

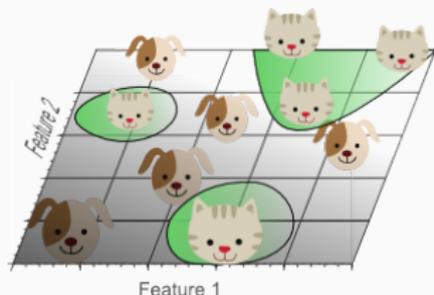
iBorderCtrl:

- iBorderCtrl asume que entre diferentes personas, etnias, géneros, edad, diversidad funcional, diversidad neuronal... hay una forma **universal** de expresar el engaño a través de lenguaje no verbal ('biomarcadores de mentiras').
- **Micro-expresiones:** expresiones faciales rápidas (parpadeo ojo izquierdo) no detectables por el ojo humano que reflejan emociones.
- Individualmente las micro-expresiones no revelan nada, pero analizando todas a la vez se pueden clasificar emociones. ¡gracias a la IA!

Problemática de la validación de iBorderCtrl (I)

Diseño experimental:

- Micro-expresiones: 1/25s-1/5s → 15fps y 30fps
- Ajuste del modelo con 86,586 patrones → 32 participantes altamente muestreados (no i.i.d.)
- 32 patrones y 38 dimensiones → maldición de la dimensión



Fuente [Computer vision for dummies](#)

Problemática de la validación de iBorderCtrl (II)

¿Cómo crear una base de datos realista?



Enjoyment



Anger



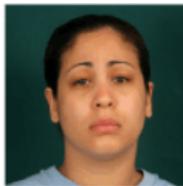
Fear



Contempt



Disgust



Sadness

PaulEkmanGroup

Problemática de la validación de iBorderCtrl (III)

relative from the EC.

- Participants are stratified into 1 of 4 blocks. These blocks are intended to have different degrees of emotional intensity / deceptive stakes (in decreasing order of intensity).
 - *S2: Simulated biohazard infectious disease in test tube with informational video about weaponization*
 - *S3: Simulated biohazard infectious disease in test tube without informational video*
 - *S4: Simulated Drug package (soap powder in clear packet)*
 - *S5: Simulated Forbidden agriculture/ food product i.e. seeds.*

Fuente [OrShea et al. \[2018\]](#)

Comentario al estilo *Cachitos de Hierro y Cromo*: Con 4.5 millones de euros deberían haber podido usar droga de verdad.

Los límites estadísticos de los tests masivos (I)

La probabilidad de que haya un persona migrante/viajera mentirosa es del 1 %.

La probabilidad de que el detector de mentiras detecte una mentira es del 73 %.

La probabilidad de que el detector de mentiras no detecte una mentira es del 24 %.

Los límites estadísticos de los tests masivos (I)

La probabilidad de que haya un persona migrante/viajera mentirosa es del 1 %.

La probabilidad de que el detector de mentiras detecte una mentira es del 73 %.

La probabilidad de que el detector de mentiras no detecte una mentira es del 24 %.

Ejercicio de estadística

Si el detector de mentiras dice que un migrante/viajero miente, ¿cuál es la probabilidad de que haya mentido? a) 0-30 % b) 30-60 % c) 60-100 %

Los límites estadísticos de los tests masivos (I)

La probabilidad de que haya un persona migrante/viajera mentirosa es del 1 %.

La probabilidad de que el detector de mentiras detecte una mentira es del 73 %.

La probabilidad de que el detector de mentiras no detecte una mentira es del 24 %.

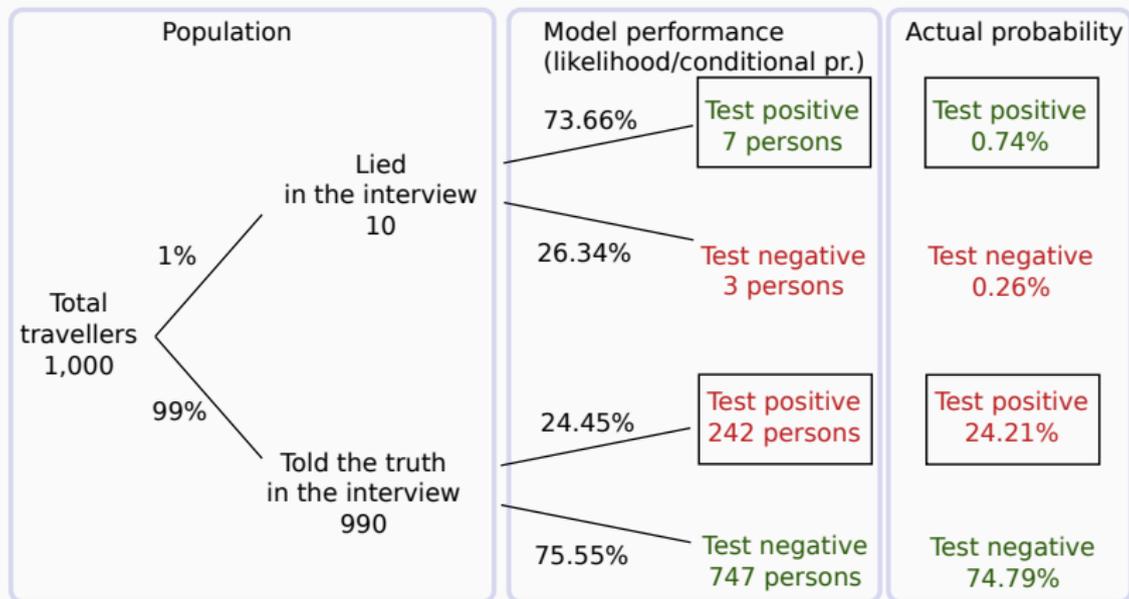
Ejercicio de estadística

Si el detector de mentiras dice que un migrante/viajero miente, ¿cuál es la probabilidad de que haya mentido? a) 0-30 % b) 30-60 % c) 60-100 %

2.9 %

Más en → [La falacia de la frecuencia base](#) y en este [cuaderno](#)

Los límites estadísticos de los tests masivos (II)



No funciona, ¿entonces?

Conclusiones:

- Es muy improbable que el detector de mentiras funcione.

No funciona, ¿entonces?

Conclusiones:

- Es muy improbable que el detector de mentiras funcione.
- Lo que si funciona es la creación de sujetos (migrantes mentirosos) a través de la tecnología y los datos.

No funciona, ¿entonces?

Conclusiones:

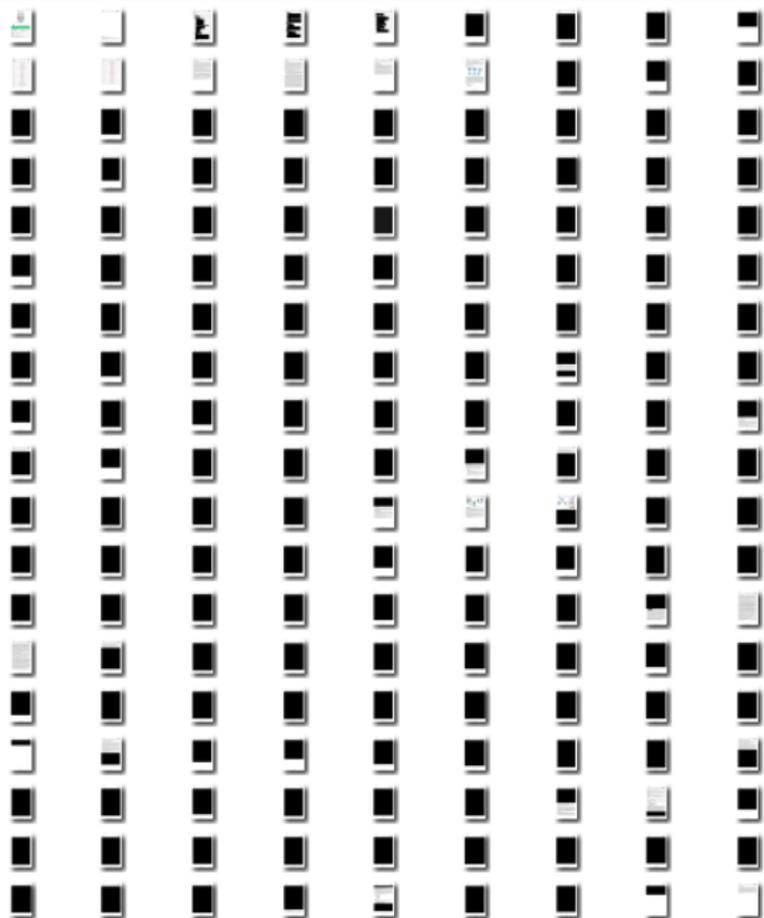
- Es muy improbable que el detector de mentiras funcione.
- Lo que si funciona es la creación de sujetos (migrantes mentirosos) a través de la tecnología y los datos.
- La función de muchas de estas herramientas y sistemas defectuosos es más bien política, y forma parte de un modelo de gobernanza.



Project deliverables

- D1.1 Ethics advisor's first report ([Confidential](#), already submitted to the European Commission)
- D1.2 Ethics of profiling, the risk of stigmatization of individuals and mitigation plan ([Confidential](#), already submitted to the European Commission)
- D1.3 Ethics Advisor ([Confidential](#), already submitted to the European Commission)
- D2.1 Requirement Analysis Report ([Confidential](#), already submitted to the European Commission)
- D2.2 Reference Architecture and components specifications ([Confidential](#), already submitted to the European Commission)
- D2.3 EU wide legal and ethical review report ([Confidential](#), already submitted to the European Commission)
- D3.1 Data Collection Devices - specifications ([Confidential](#), already submitted to the European Commission)
- D3.2 First version of all technological tools and subsystems (Portable unit, ADDS, DAAT, portable radar, FMT, Avatar based dialogue) ([Confidential](#), already submitted to the European Commission)
- D3.3 Second version of all technological tools and subsystems for integration ([Confidential](#), already submitted to the European Commission)
- D4.1 First version of the iBorderCtrl software platform ([Confidential](#), already submitted to the European Commission)
- D4.2 Second version of the iBorderCtrl software platform ([Confidential](#), already submitted to the European Commission)
- D5.1 Integration Plan ([Confidential](#), already submitted to the European Commission)
- D5.2 Early version of the integrated prototype (limited functionality) ([Confidential](#), already submitted to the European Commission)
- D6.1 Experimental Design for Pilot Deployment and Evaluation ([Confidential](#), already submitted to the European Commission)
- D7.1 Project Web Portal

iBorderCtrl: FOI request



Conclusiones y debate

Javier Sánchez-Monedero and Lina Dencik. The politics of deceptive borders: 'biomarkers of deceit' and the case of iBorderCtrl. *Information, Communication & Society*. 2020.

<https://doi.org/10.1080/1369118X.2020.1792530>

<https://datajusticeproject.net/resources/>

Llamémosle mierda.



Llamémosle mierda.



<https://www.callingbullshit.org/>

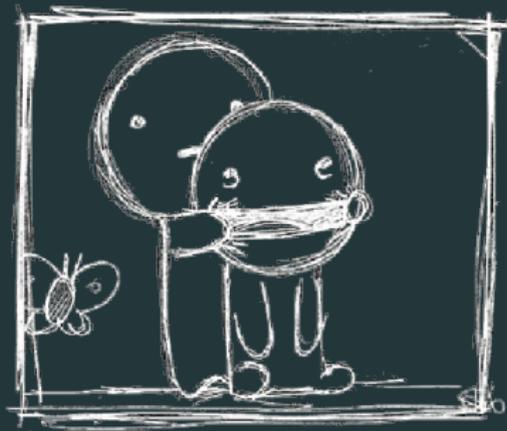
Conclusiones

Problemática general de la **inteligencia artificial emocional** / **computación afectiva**:

- Marco reduccionista forzoso tanto categórico como contextual
- Imposibilidad de crear escenarios de validación realistas
- Riesgo de descubrimiento (¿y creación?) de proxies hacia variables protegidas
- Las narrativas sobre IA están trayendo de vuelta la frenología
- 6 emociones en la cara...¿y por qué 6 en el lenguaje?
- Los proyectos para cuantificar el riesgo a través de la evaluación del comportamiento y emociones están en **auge**



Gràcies!



- L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez, and S. D. Pollak. Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychological Science in the Public Interest, 20(1):1–68, July 2019. ISSN 1529-1006, 1539-6053. doi: 10.1177/1529100619832930. URL <http://journals.sagepub.com/doi/10.1177/1529100619832930>.
- National Research Council. The Polygraph and Lie Detection. The National Academies Press, Washington, DC, 2003. ISBN 978-0-309-26392-4. doi: 10.17226/10420. URL <https://www.nap.edu/catalog/10420/the-polygraph-and-lie-detection>.
- J. OrShea, K. Crockett, W. Khan, P. Kindynis, A. Antoniadis, and G. Bouladakis. Intelligent Deception Detection through Machine Based Interviewing. In 2018 International Joint Conference on Neural Networks (IJCNN), pages 1–8, Rio de Janeiro, Brazil, July 2018. IEEE. ISBN 978-1-5090-6014-6. doi: 10.1109/IJCNN.2018.8489392. URL <https://ieeexplore.ieee.org/document/8489392/>.