



Datos en el contexto de la salud, del bienestar y de la autonomía personal

Jordi Conesa i Caralt



+



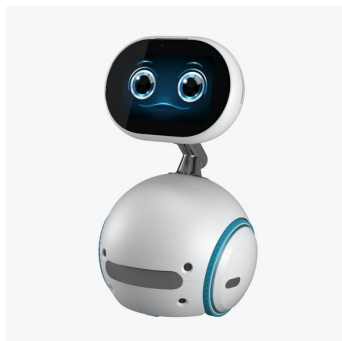
+



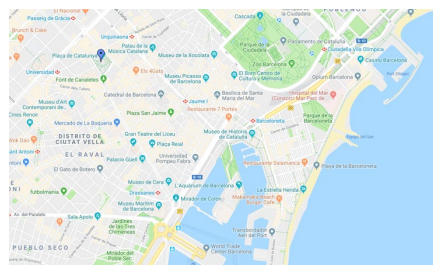
+



+



+



+

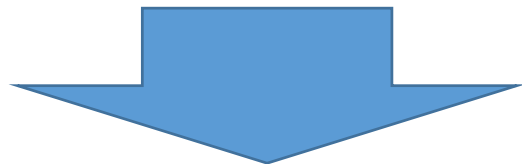


+



Big Data: ¿Qué es?

Conjunto de **técnicas y tecnologías** para el **tratamiento y almacenamiento de datos**, en entornos de **gran volumen**, **variedad** de orígenes y en los que la **velocidad** de respuesta es crítica



Necesitamos un conjunto de computadores trabajando colaborativamente

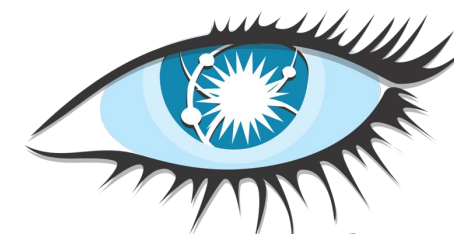


I Congreso sobre el derecho a la autonomía personal

DISCAPACIDAD FÍSICA Y ORGÁNICA, ENVEJECIMIENTO Y CRONICIDAD



elasticsearch

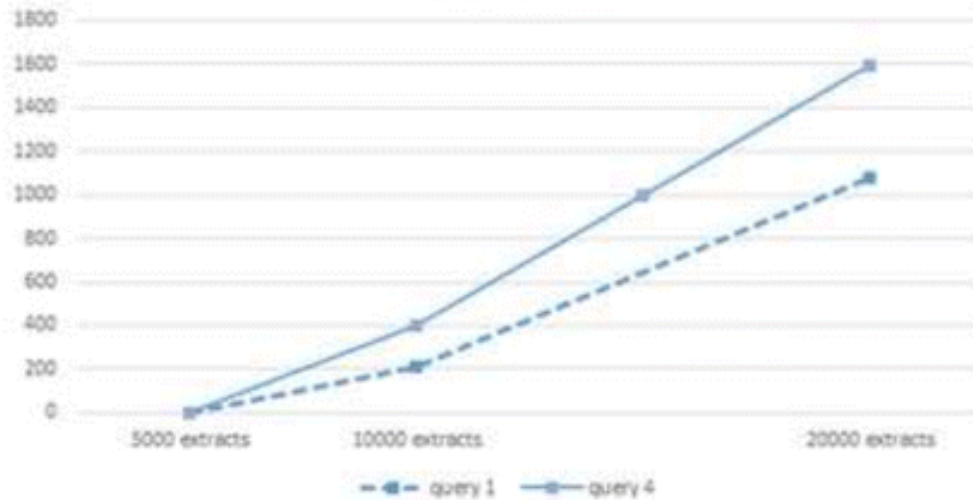


cassandra

Ver <https://vimeo.com/223539974>

Almacenamiento y recuperación de EHR

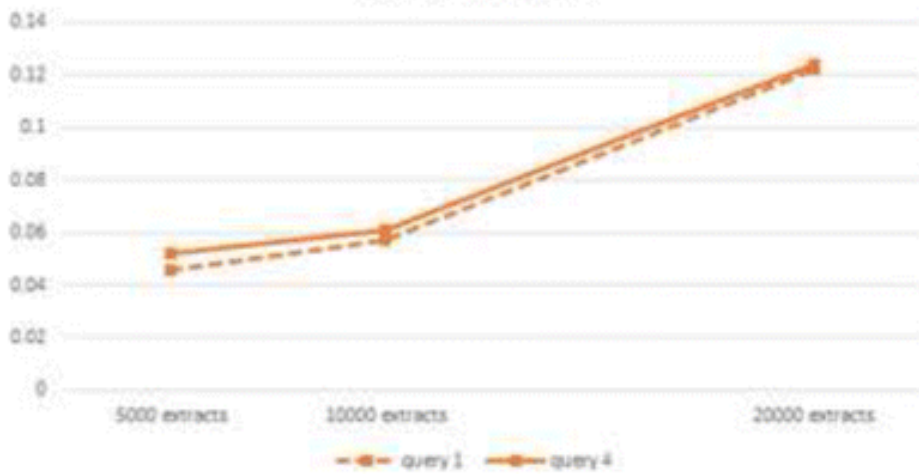
Q1, Q4; ORM



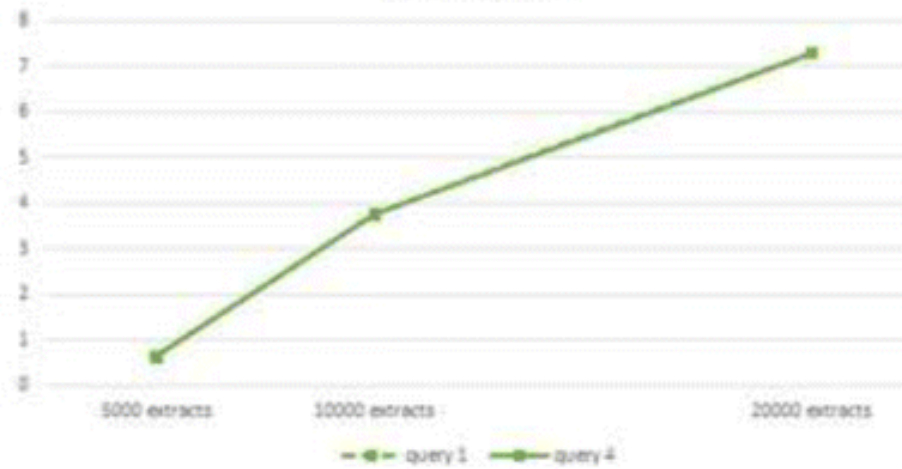
Q1 Find all problems of a single patient

Q4 Find initial date, resolution date and severity of all problems of a single patient

Q1, Q4; MongoDB

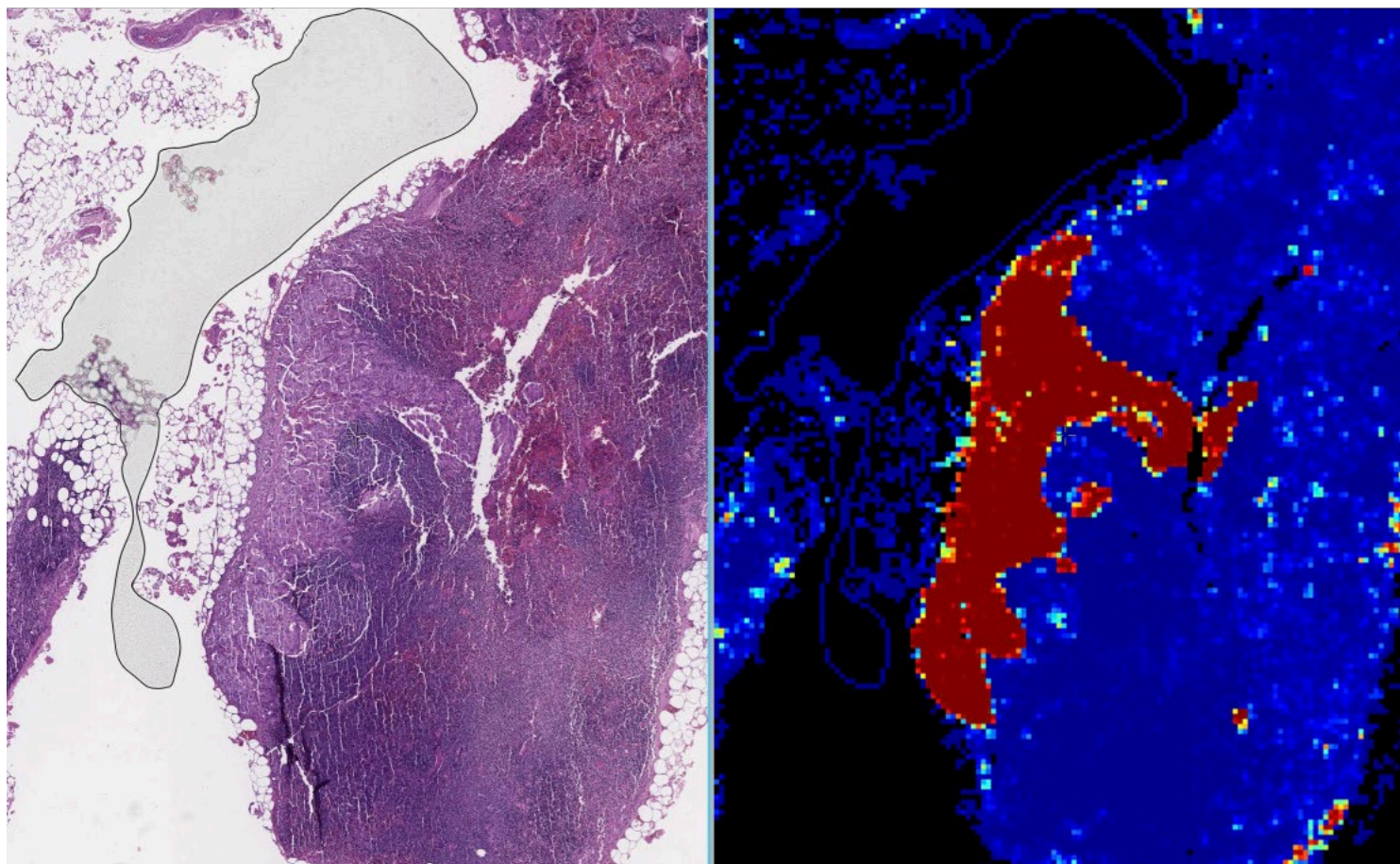


Q1, Q4; eXist



Fuente: Sánchez-de-Madariaga, Ricardo, et al. "Examining database persistence of ISO/EN 13606 standardized electronic health record extracts: relational vs. NoSQL approaches." *BMC medical informatics and decision making* 17.1 (2017): 123.

Sistemas predictivos muy precisos



Fuente: Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., ... & Hipp, J. D. (2017). Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv:1703.02442*.

ENTER PATIENT'S DATA

Age	78	Black Race	No	Cardio-vascular disease (clinical or subclinical)	No
Sex	Female	Smoking Status	Former Smoker	eGFR (mL/min/1.73 m ²)	47
Weight (Kg)	77	Systolic blood pressure (mmHg)	155	Total cholesterol (Mg/dL)	212
Height (M)	1.63	Blood pressure lowering medications (number)	0	High density cholesterol (HDL) (Mg/dL)	75

Integrar las opiniones del paciente en los sistemas inteligentes



CALCULATE RISK

RESULTS

Severity rank	Cardiovascular improvement due to intensive treatment*	Severity rank	Adverse events due to intensive treatment*
---------------	--	---------------	--

Acute Myocardial Infarction (iNNT=384)

1.4% -0.3%

Serious Hypotension (iNNH=61)

2.5% +1.6%

Acute Decompensated Heart Failure (iNNT=113)

2.2% -0.9%

Serious Syncope (iNNH=75)

4.5% +1.3%

Stroke (iNNT=442)

1.9% -0.2%

Serious Electrolyte Abnormality (iNNH=47)

6.4% +2.1%

Cardio-Vascular Death (iNNT=101)

2.3% -1.0%

Serious Acute Renal Failure (iNNH=32)

3.0% +3.1%

RECOMMENDATION

Do Not Treat BP Intensively

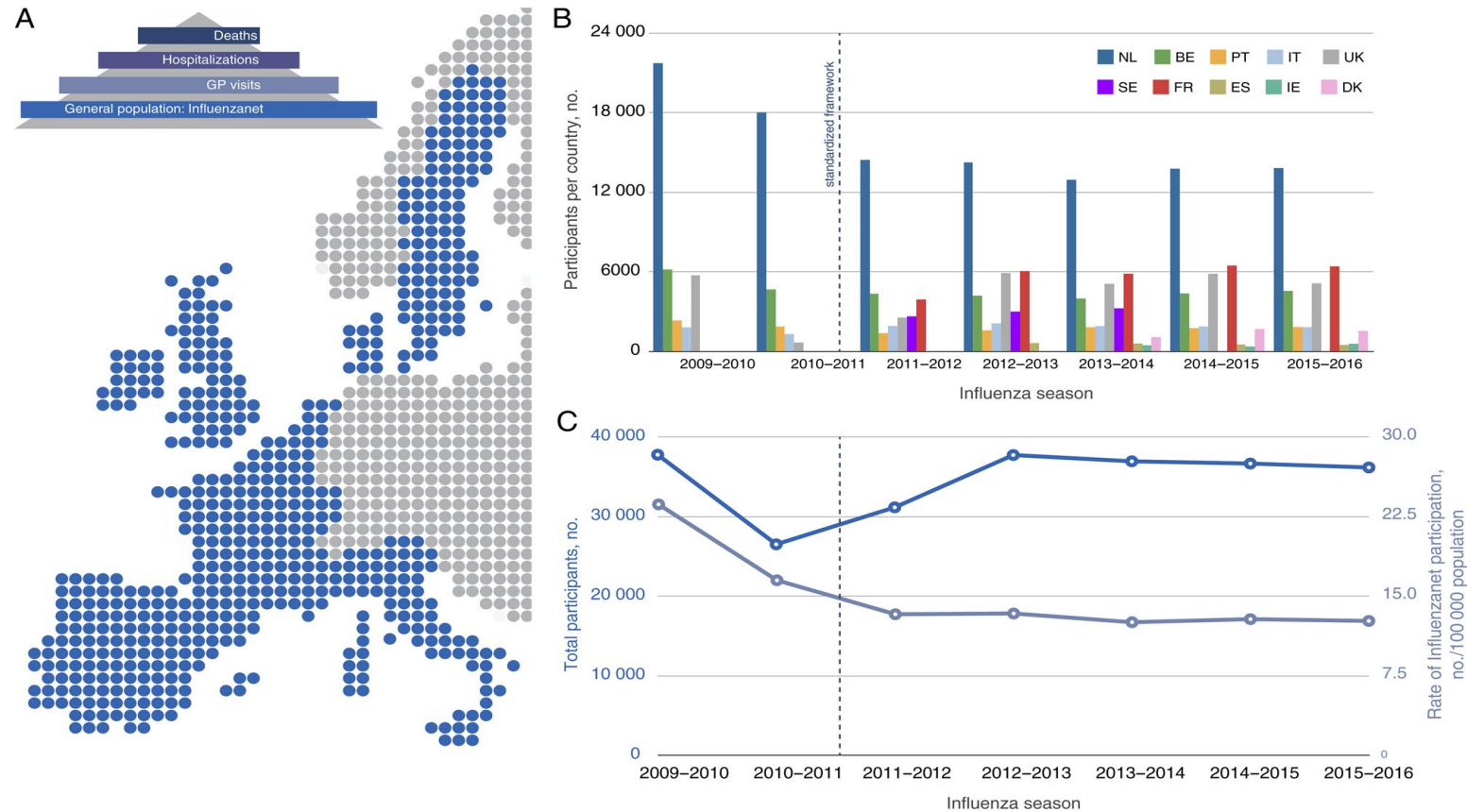
Systolic blood-pressure target:
≤140 mmHg

The recommendation is based on the ratio between the individual ARRs and ARIs, weighted by severity ranks assigned to the different outcomes.

The current ranks are averages of ranks given by several physicians. You can change the ranks and update the recommendation.

UPDATE RECOMMENDATION
with new severity ranks

El Big Data no es solo máquinas y análisis de datos

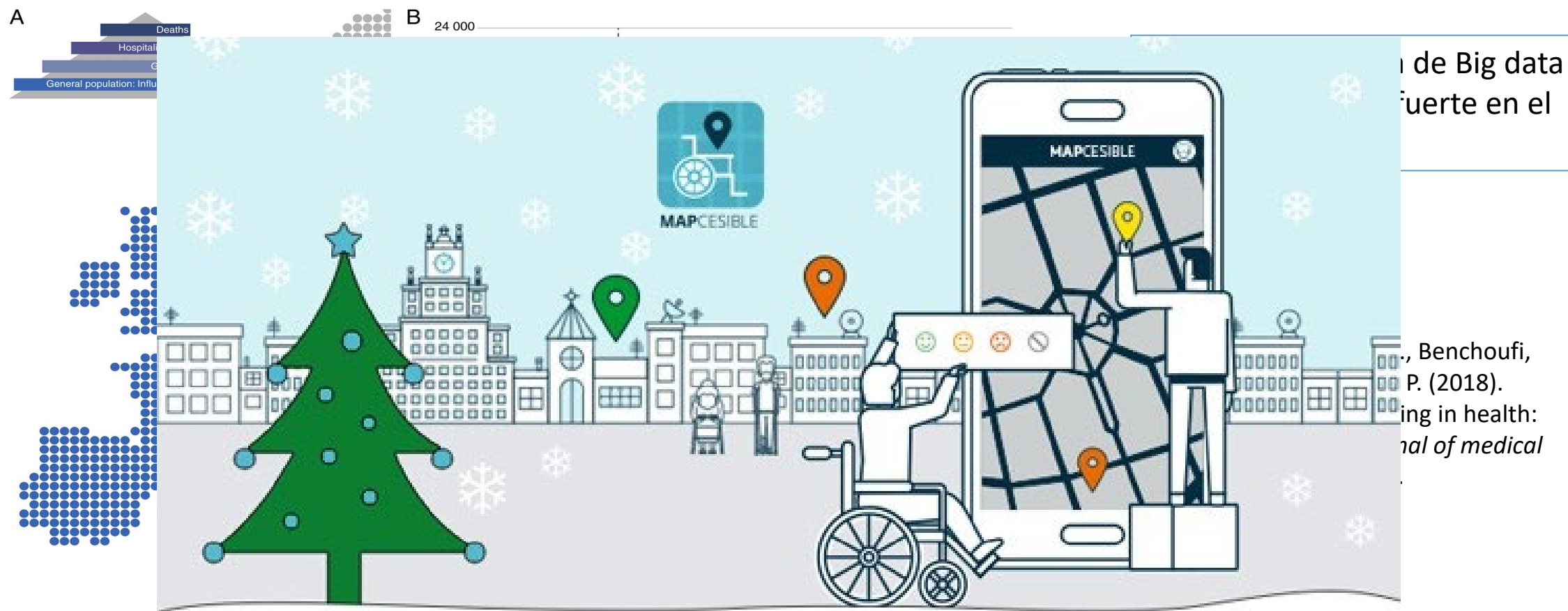


Una nueva filosofía de Big data
que está pegando fuerte en el
sector salud

Créquit, P., Mansouri, G., Benchoufi, M., Vivot, A., & Ravaut, P. (2018). Mapping of crowdsourcing in health: Systematic review. *Journal of medical Internet research*, 20(5).

Fuente: Koppeschaar, C. E., Colizza, V., Guerrisi, C., Turbelin, C., Duggan, J., Edmunds, W. J., ... & Paolotti, D. (2017). Influenzanet: citizens among 10 countries collaborating to monitor influenza in Europe. *JMIR public health and surveillance*, 3(3).

El Big Data no es solo máquinas y análisis de datos



Fuente: Koppeschaar, C. E., Colizza, V., Guerrisi, C., Turbelin, C., Duggan, J., Edmunds, W. J., ... & Paolotti, D. (2017). Influenzanet: citizens among 10 countries collaborating to monitor influenza in Europe. *JMIR public health and surveillance*, 3(3).

¿Y en el ámbito de la autonomía personal?

- *Smart Homes*
- *Smart Wheelchairs*
- *Smart Cushions*
- Sistemas de ayuda (*Self-help devices*)
- *Smart glasses*
- *Smart piercings*
- Sistemas de navegación autónomos (o semi-autónomos)
- Control de dispositivos mediante ondas cerebrales
- Exoesqueletos, ...

Fuentes de interés:

- Leaman, J., & La, H. M. (2017). A comprehensive review of smart wheelchairs: past, present, and future. *IEEE Transactions on Human-Machine Systems*, 47(4), 486-499.
- Orejuela-Zapata, J. F., Rodríguez, S., & Ramírez, G. L. (2019). Self-Help Devices for Quadriplegic Population: A Systematic Literature Review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4), 692-701.

¿Cuál es nuestro grado de madurez?

Hemos avanzado mucho

Pero aún queda un largo camino que recorrer



I Congreso sobre el derecho a la autonomía personal
DISCAPACIDAD FÍSICA Y ORGÁNICA, ENVEJECIMIENTO Y CRONICIDAD

Es nuestra
primera
oportunidad
de obtener
casi el 100% de
la información
relevante para
la salud



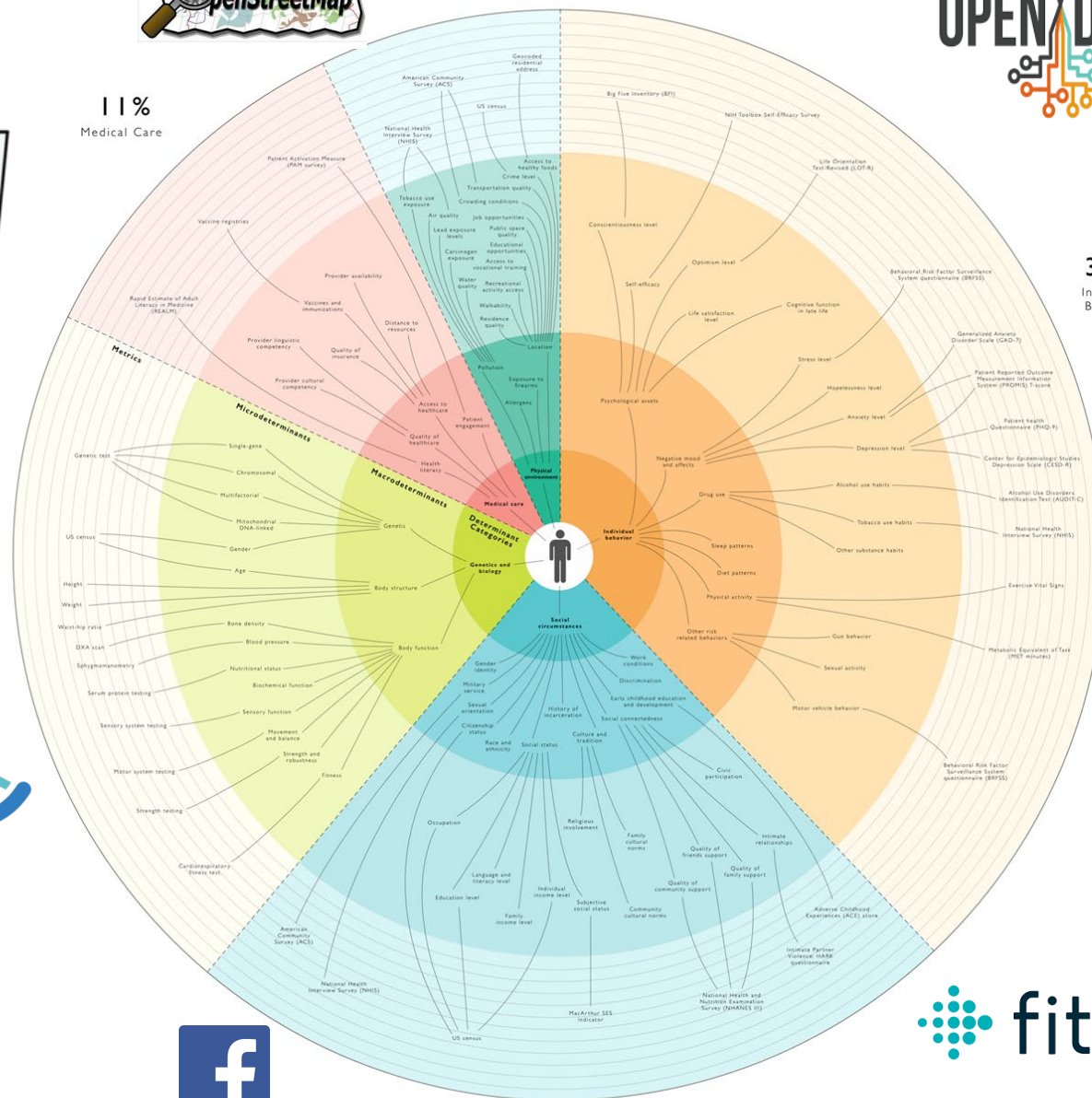
7%
Physical
Environment



11%
Medical Care

22%
Genetics and
Biology

36%
Individual
Behavior

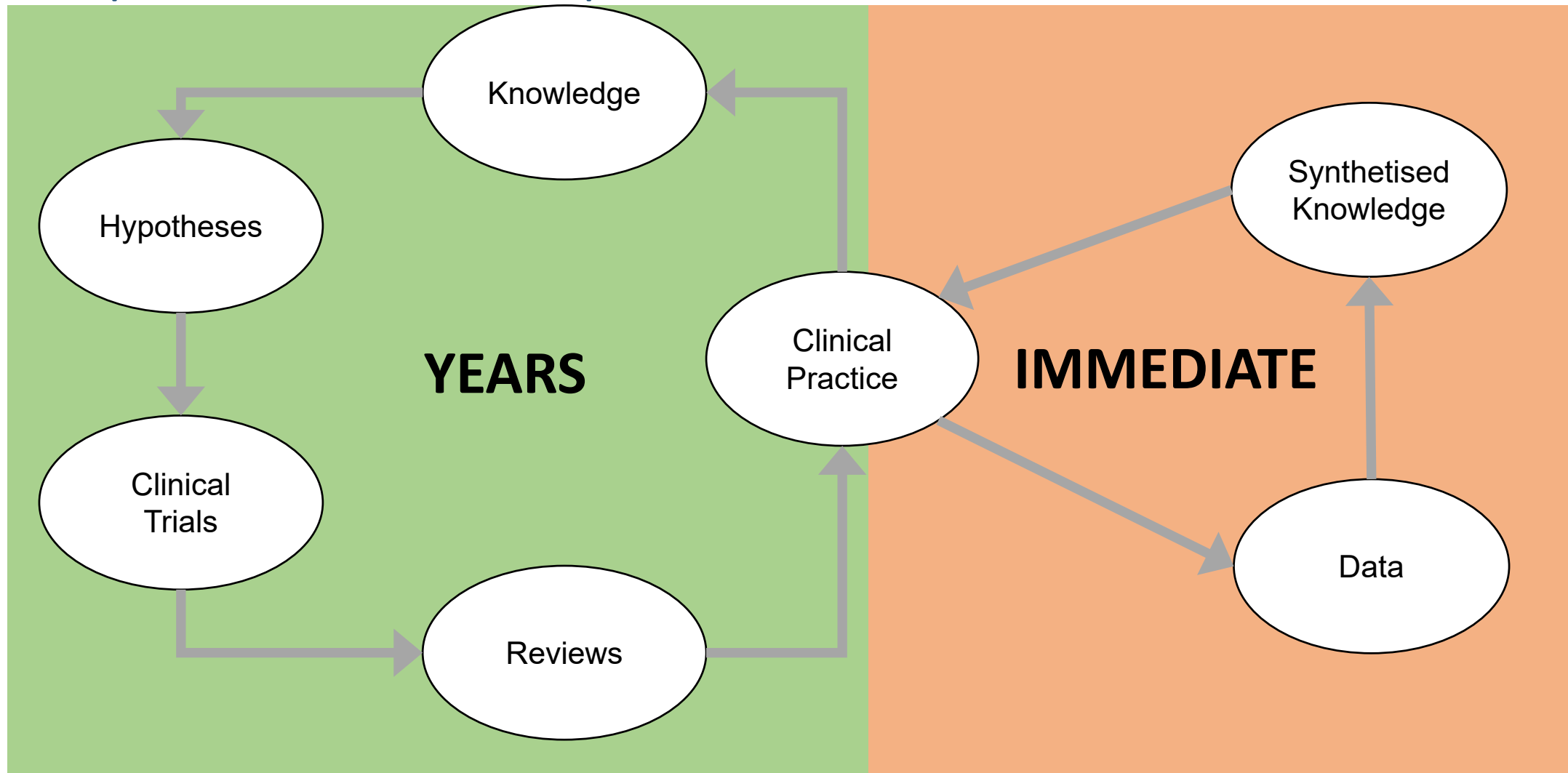


24%

Social Circumstances



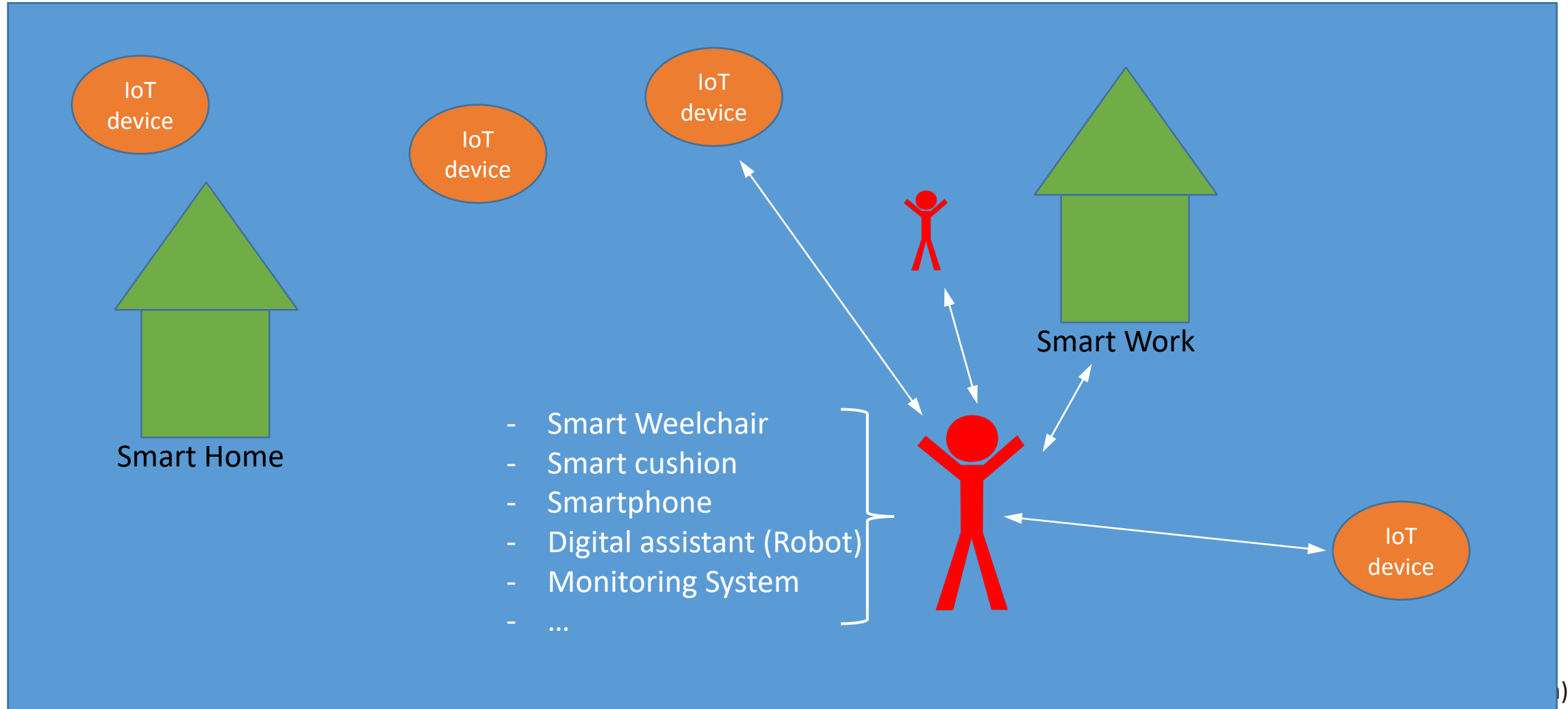
Y de aprender más rápido



Fuente: Stein Olav Skrøseth. Health data science. Building Learning Healthcare Systems from Big Data: Challenges and Promises. 1st International Research Conference on eHealth and eWell-being (2018).

Y de promover la autonomía – ecosistema *smart*

Smart Cities



Vamos a hablar de datos y de...

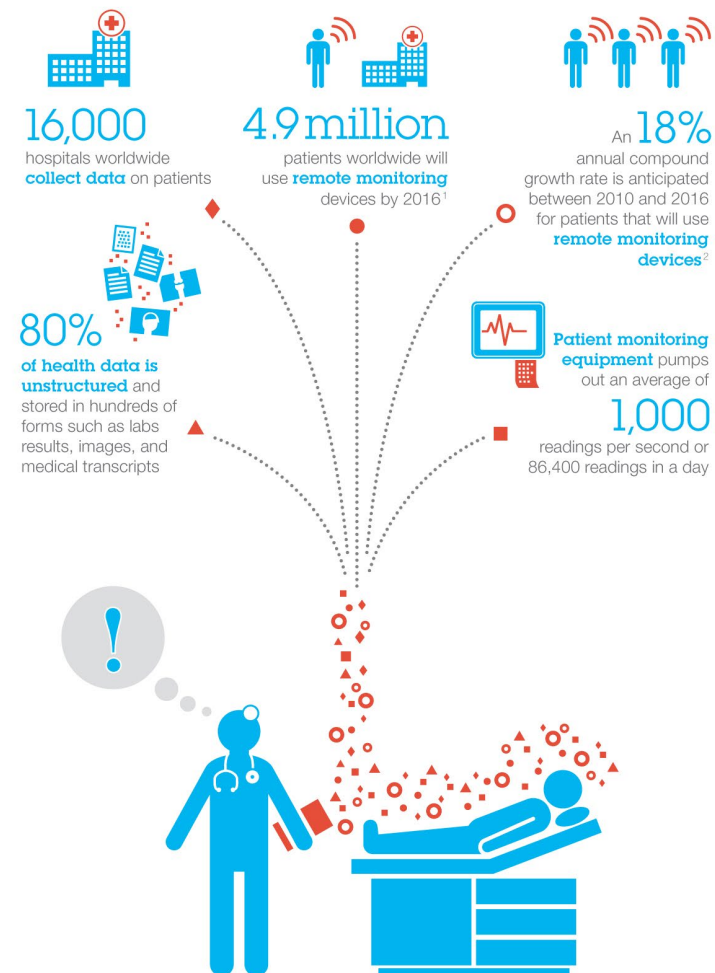
- Algunos ejemplos,
- La oportunidad actual,
- Los **retos** a considerar,
- Y de las tecnologías disponibles.

Hay multitud de datos, pero ¿donde?

Big Data in Healthcare: Tapping New Insight to Save Lives

COCEMFE

Healthcare is challenged by large amounts of data in motion that is diverse, unstructured and growing exponentially. Data constantly streams in through interconnected sensors, monitors and instruments in real-time faster than a physician or nurse can keep up.



The ability to analyze big data in motion in real-time as it streams in can help predict the onset of illness and respond instantly from new insight that will help transform healthcare.

Compartir datos tiene un gran impacto

Encuesta sobre discapacidades, autonomía personal y situaciones de dependencia
2008

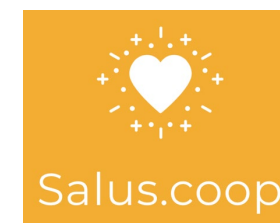
- Según Google Académico se han generado **más de 1450 artículos científicos** al respecto
- Aún teniendo en cuenta las limitaciones de los datos:
 - No tienen información longitudinal
 - Recogida mediante encuestas: problemas de subjetividad, fiabilidad, etc.

¿Cómo promover el acceso a los datos?

- Mecanismos para integrar datos (médicos y no médicos)
- Mecanismos para compartir datos

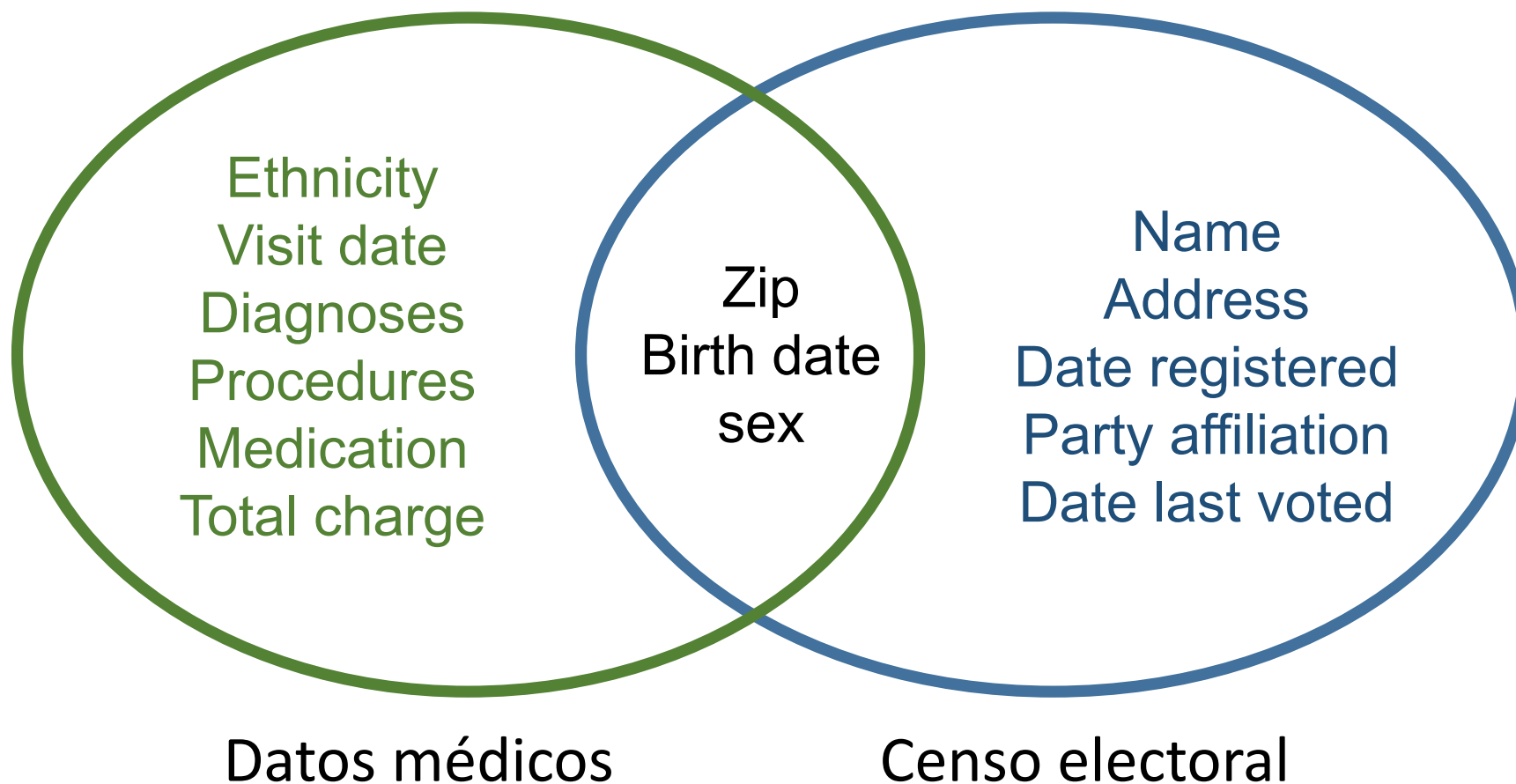


- Mecanismos que permitan empoderar los pacientes en la creación y gestión de TODOS sus datos de salud



- Al compartir datos aparecen los problemas de privacidad

87% de los americanos se pueden identificar por el código postal, la fecha de nacimiento y el sexo



Privacidad de datos

- No importa como se anonimicen los datos, el riesgo de re-identificación siempre existe.
- Algunos atributos (código postal, IP, ID del teléfono, geo posición...) pueden llevar a identificarnos fácilmente.
 - En algunos trabajos (y países), el código postal y la información han llevado a detectar la raza, sexo e incluso religión de gran parte de la población



- No obstante en el ámbito nacional se está haciendo un gran trabajo en el tema y las garantías son máximas

Los algoritmos de *machine learning* discriminan si no lo evitamos (sesgo algorítmico)



Discriminación y Big Data

- Los algoritmos de ML tienden a exacerbar las desigualdades de los datos
 - Hay muchos casos en entornos de alojamiento, empleabilidad, crédito, consumo, etc.
 - Entre 2004 y 2009 la comunidad afroamericana y hispana de EUA pagaron cerca de un 30% más en sus hipotecas
- Hay cierta indefensión de los generadores de datos respecto a los recolectores y analistas
 - Muchas veces no tiene claro qué datos se están recogiendo y para qué
- Sería conveniente promover la transparencia: mostrar **qué datos** tienen y **qué inferencias** han hecho con ellos: “potencial diabético”, “adulto con padre anciano”, ...
- Hay técnicas para eliminar el sesgo y hacer los algoritmos más justos

Discriminación y Big Data

- Tema candente en el contexto de investigación
 - En una revisión sistemática del 2019, 49 de los 61 artículos eran de los últimos 5 años

Fuente: Favaretto, M., De Clercq, E., & Elger, B. S. (2019). Big Data and discrimination: perils, promises and solutions. A systematic review. *Journal of Big Data*, 6(1), 12.

Otros retos detectados

- Algunos algoritmos de interacción con las personas tienen margen de mejora: interface cerebral, reconocimiento de la habla, integración modal de diversas tecnologías, problemas en entornos indoor, etc.
- Los sistemas de asistencia:
 - No deben ser standard, deben personalizarse para cada usuario.
 - Deben ser multimodales, hasta ahora no lo son.
 - Son demasiado costosos
 - Tienen una adherencia muy baja (30% abandona después de su primer uso).
- La calidad de los conjuntos de datos de discapacidad son escasos, mejorables y requieren de preprocesado antes de ser analizados

¿Y la tecnología es un reto pendiente?

No, la tecnología está lista.

Big Data

Machine Learning

Block Chain



Hemos avanzado mucho, pero aún nos queda un largo camino para recorrer, pero el destino lo merece. Usemos la tecnología y el uso adecuado de los datos para promover nuestro bienestar y nuestra autonomía personal.

Jordi Conesa i Caralt

Coordinador del ámbito de Ciencia de Datos en Salud del
eHealth Center

jconesac@uoc.edu

jconesac.wordpress.com/

@Jordi_Conesa

