

Business Value of Privacy in Tourism Data Spaces

- enabled by PET -



CMMIDEV/3SM
Exp. 2019-04-22 / Appraisal #26155

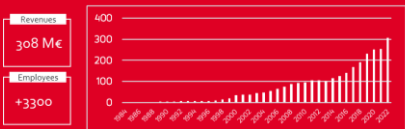
GMV SOLUCIONES GLOBALES INTERNET S.A.U.
© GMV 2024 Todos los derechos reservados

El presente documento está clasificado en nivel "GMV-ICL2-Limited". Esta clasificación habilita a su receptor al uso de la información contenida en el documento para los fines para los que la empresa la ha facilitado o, en su caso, a lo acordado contractualmente en relación al intercambio de información entre las partes, y ello sin perjuicio del cumplimiento de la normativa sobre propiedad intelectual y sobre protección de datos de carácter personal.



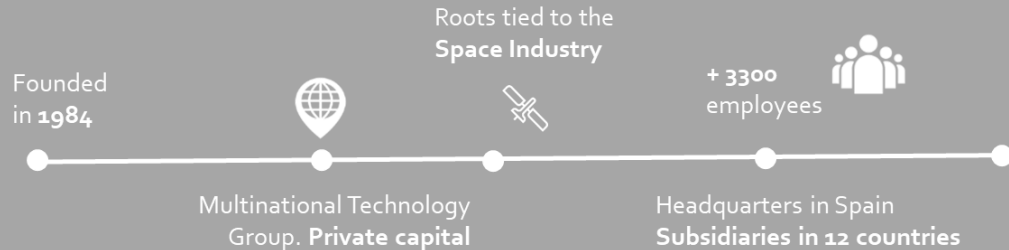
Technology Leadership

GMV in figures: sustained growth

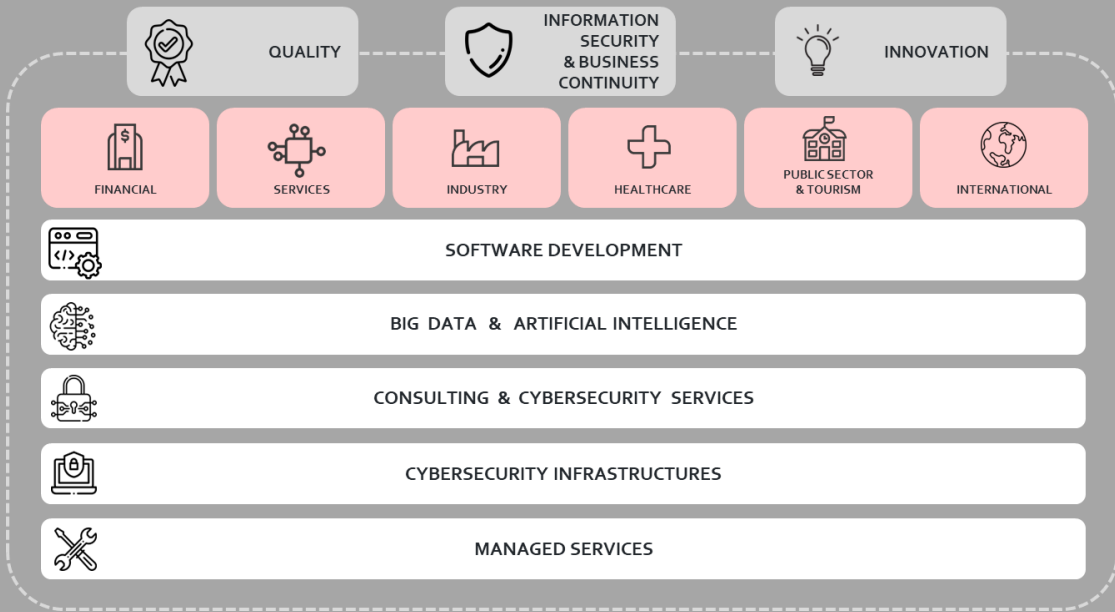


GMV in the World

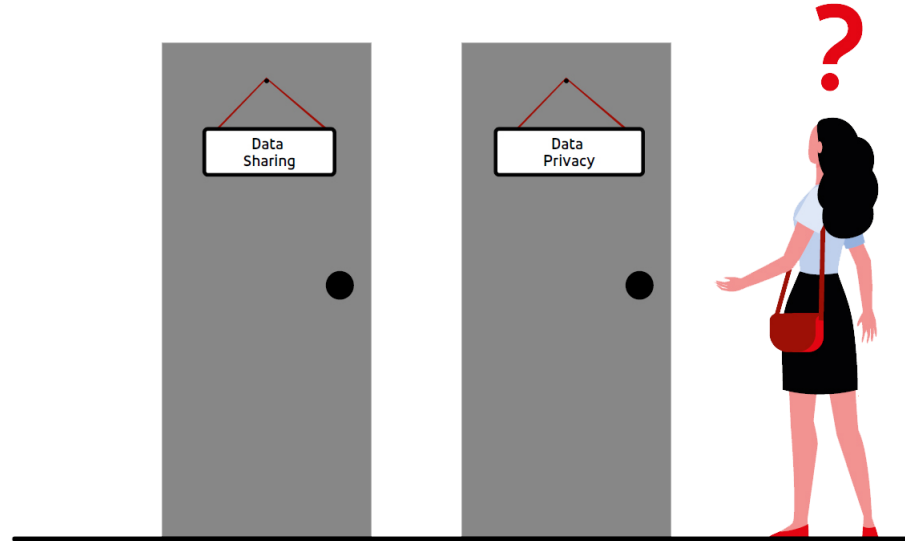
- SPAIN**
 - Madrid
 - Valladolid
 - Sevilla
 - Barcelona
 - Valencia
 - Zaragoza
- INTERNATIONAL**
 - Colombia
 - France
 - Germany
 - Malaysia
 - USA
 - Portugal
 - Poland
 - Romania
 - United Kingdom
 - Belgium
 - The Netherlands



Engineering, development and integration of systems, software, hardware, specialized products and services



¿Es insuperable el equilibrio entre la **privacidad** y el **intercambio** de los datos?





Framework de computación distribuida desarrollado por GMV



Permite realizar cálculos de forma segura y privada sobre los datos distribuidos sin exponerlos ni moverlos



Aprovecha los métodos criptográficos avanzados que mantienen los datos cifrados mientras se procesan durante el cómputo



Compartir de forma segura
el conocimiento basado en
sus datos



Construir y mejorar modelos de
machine learning cumpliendo con
la privacidad de las fuentes de
datos distribuidas



Facilitar el intercambio
seguro de conocimiento
entre organizaciones

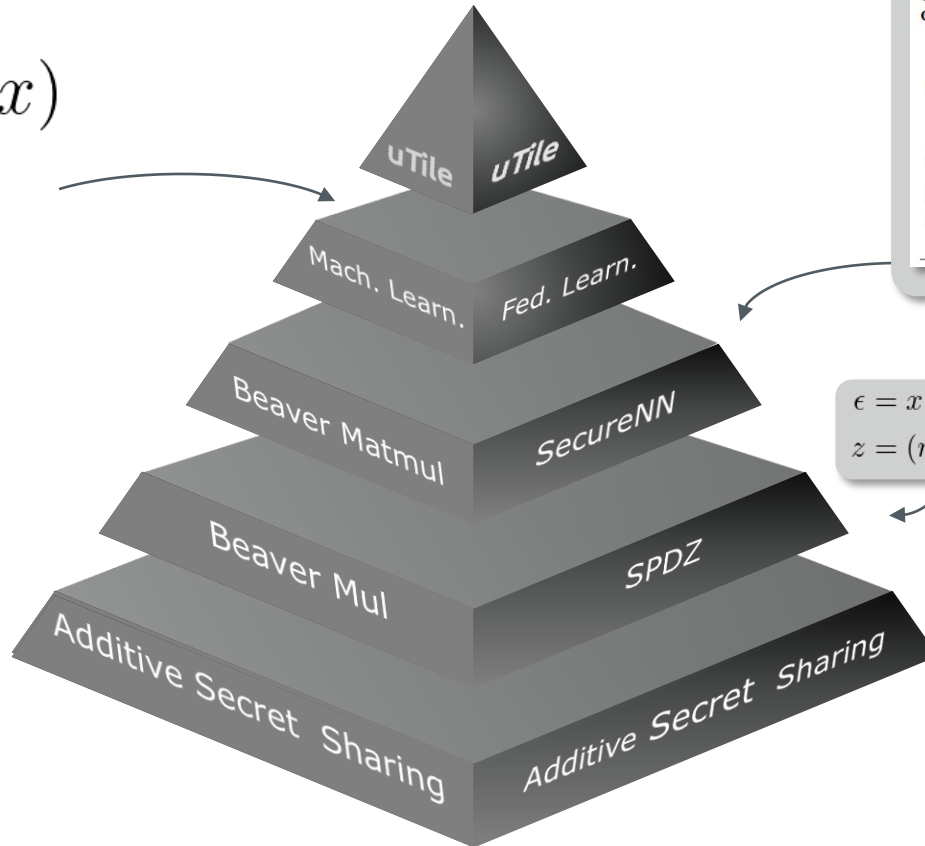


$$\operatorname{argmax}_{x \in S} f(x)$$

$$Z = X \times Y$$

$$z = x \cdot y$$

$$z = x + y$$



Algorithm 1 Mat. Mul. $\Pi_{\text{MatMul}}(\{P_0, P_1, P_2\})$:

Input: P_0 & P_1 hold $(\langle X \rangle_0^L, \langle Y \rangle_0^L)$ & $(\langle X \rangle_1^L, \langle Y \rangle_1^L)$ resp.

Output: P_0 gets $\langle X \cdot Y \rangle_0^L$ and P_1 gets $\langle X \cdot Y \rangle_1^L$.

Common Randomness: P_0 and P_1 hold shares of zero matrices over $\mathbb{Z}_L^{m \times v}$ resp.; i.e., P_0 holds $\langle 0^{m \times v} \rangle_0^L = U_0$ & P_1 holds $\langle 0^{m \times v} \rangle_1^L = U_1$

- 1: P_2 picks random matrices $A \xleftarrow{\$} \mathbb{Z}_L^{m \times n}$ and $B \xleftarrow{\$} \mathbb{Z}_L^{n \times v}$ and generates for $j \in \{0, 1\}$, $\langle A \rangle_j^L, \langle B \rangle_j^L, \langle C \rangle_j^L$ and sends to P_j , where $C = A \cdot B$.
- 2: For $j \in \{0, 1\}$, P_j computes $\langle E \rangle_j^L = \langle X \rangle_j^L - \langle A \rangle_j^L$ and $\langle F \rangle_j^L = \langle Y \rangle_j^L - \langle B \rangle_j^L$.
- 3: P_0 & P_1 reconstruct E & F by exchanging shares.
- 4: For $j \in \{0, 1\}$, P_j outputs $-jE \cdot F + \langle X \rangle_j^L \cdot F + E \cdot \langle Y \rangle_j^L + \langle C \rangle_j^L + U_j$.

$$\epsilon = x - r_x ; \delta = y - r_y$$

$$z = (r_x \cdot r_y) + \epsilon \cdot r_y + \delta \cdot r_x + \epsilon \cdot \delta$$

$$s_n = s - \sum_{i=1}^{n-1} r_i \bmod F$$



SMPC (Secure Multi-Party Computation)

Secure Multi-party Computation makes use of additive secret sharing, which allows private data to be segmented into parts in such a way that none of the participants can reconstruct the original data, but all of them can benefit from the outcome of the sharing.

FL (Federated Learning)

Federated Learning is an algorithmic solution that allows training of ML models by sending copies of a model and the performance of training where the data are stored, thus eliminating the need to share data on a central server.

PSI (Private Set Intersection)

It is used in cases of vertical partition. PSI is a cryptographic technique that allows the intersection of data sets without having to reveal private information.

DP (Differential Privacy)

This technique allows data privacy to be maintained by deliberately injecting a degree of noise to allow statistical analysis but not re-identification. The amount of noise is controlled by the so-called privacy budget.

HOTEL A



$$40\text{€} = -\cancel{66} + \cancel{388} + 688$$

81

HOTEL B



$$60\text{€} = \color{blue}{62} - 20 + \color{blue}{118}$$

88

HOTEL C

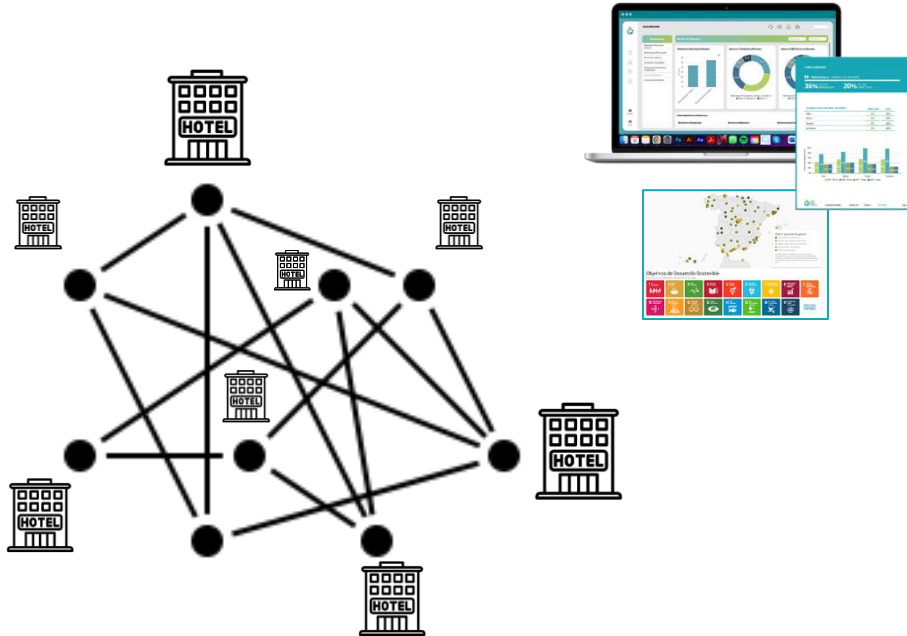


$$80\text{€} = -\color{green}{75} + -\color{green}{112} + \color{green}{167}$$

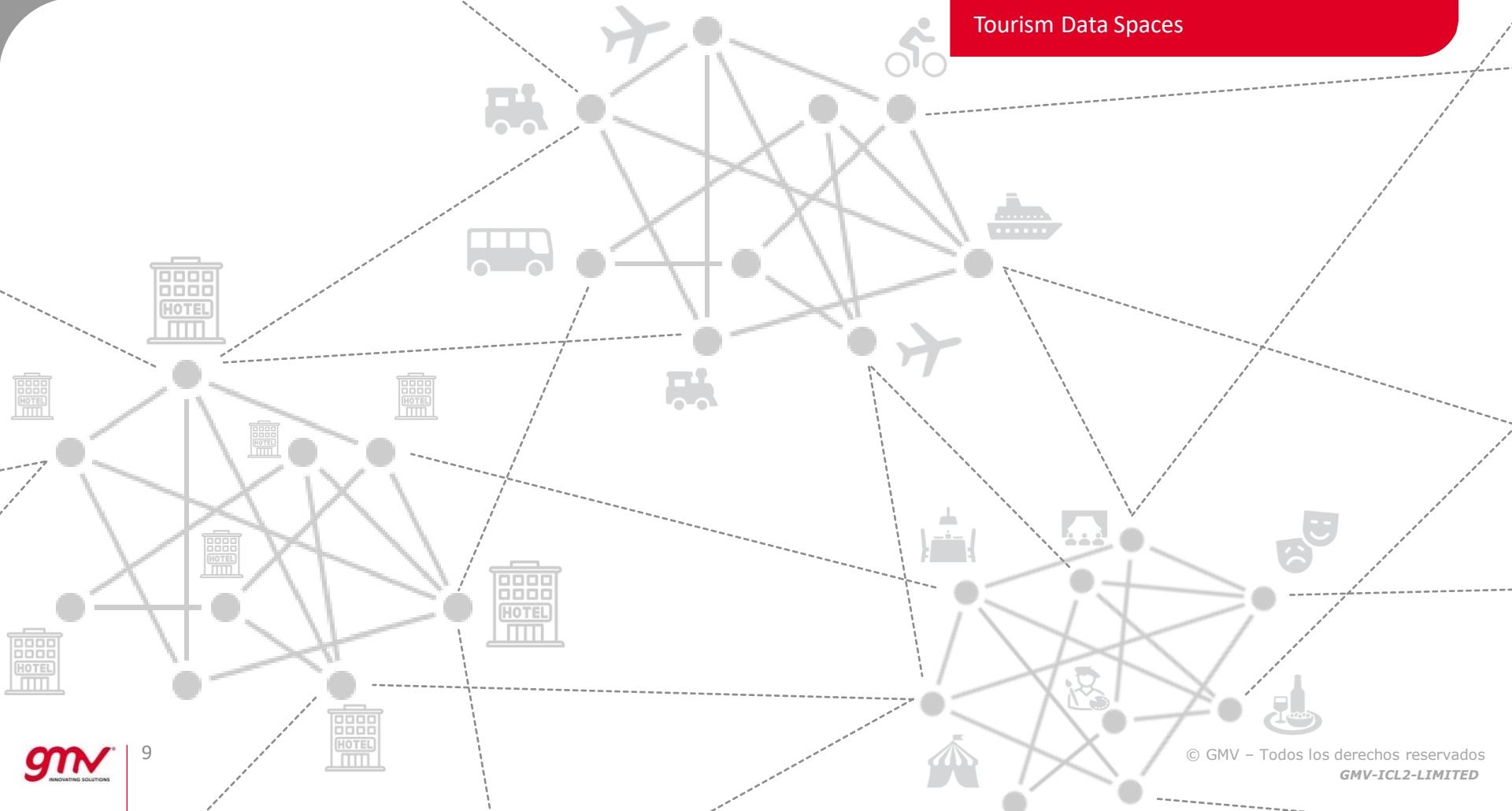
11

Resultado compartido

$$\frac{\quad + \quad}{3} = 60$$



- Revenue Management
- Calculate CO2 footprint, improve energy efficiency, measure the impact on the destination and promote compensation models
- Cybersecurity
- Sociodemographic and Semantic Analysis
- MICE (by relevance, industry, destination, human resources, sustainability, prices...)



Tourism Data Space



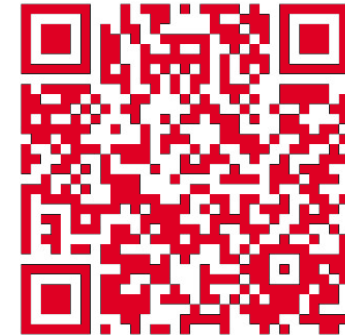
- 1 Demostrar que es factible realizar y **acelerar la investigación usando tecnología de análisis y aprendizaje federado** de una forma segura desde las Instituciones Sanitarias Españolas.
- 2 Establecer la red federada como factor clave para que todas las regiones e **instituciones de Salud Nacionales puedan contribuir y beneficiarse de los bancos de datos**, permitiendo un mejor entrenamiento de los modelos matemáticos que permitan ayudar a la toma de decisiones en el diagnóstico y tratamiento en los casos de Alzheimer, Degeneración Macular, Cáncer de Próstata, Crónico Complejo y Diagnóstico por ultrasonido.
- 3 **Pilotar la implementación de la IA en los sistemas de salud** desde el desarrollo de los algoritmos hasta la experiencia de los usuarios (pacientes / sanitarios / profesionales de tecnología) pasando por la transformación de sistemas, procesos, aprendizaje y cultura

gmv.com



Joan Antoni Malonda
Tourism Business Developer

✉ jamalonda@gmv.com



Muchas gracias