

TIC 360

ANALÍTICA DE DATOS EN LA UNIVERSIDAD

2023



| ÍNDICE

0. Introducción: Sin datos no hay paraíso	4
1. La oportunidad del concepto de Espacio del Datos	10
2. European Skill Data Space	24
3. La universidad española en el contexto del los espacios europeos de datos	34
4. Certificaciones en el ámbito de los datos	44
5. La Universidad en la era del dato	62

Versión digital



Depósito legal M-31563-2023

ISBN: 978-84-09-55266-5

Licencia



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

0. INTRODUCCIÓN: SIN DATOS NO HAY PARAÍSO

Joaquín Canca

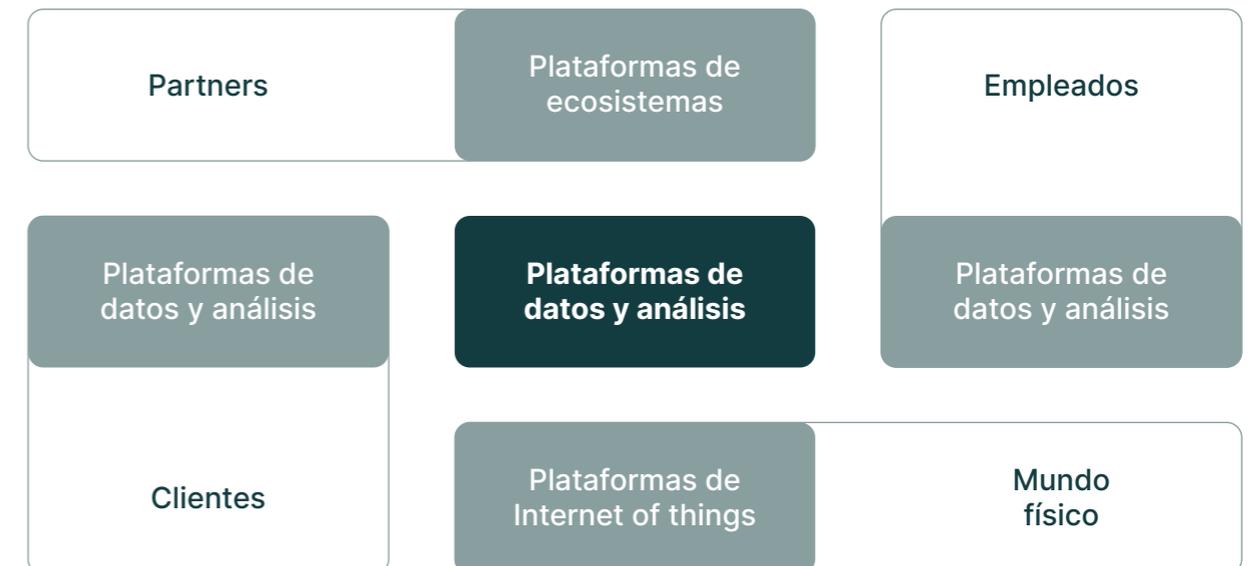
Universidad de Málaga



Analítica de datos, análisis de datos, economía del dato, gestión de datos, ciencia de los datos, minería de datos, Datawarehouse, Big Data, Data-Driven... son conceptos que se han convertido en toda una filosofía de nuestros tiempos, hasta el punto de que David Brooks, analista cultural del New York Times, acuñó el término «dataísmo» para referirse a ella. Noah Harari va más allá y lo considera como una religión que «no venera ni a dioses ni al hombre: adora los datos». Y es que pareciera que, como se titula un reciente espectáculo teatral, «Sin datos no hay paraíso».

Echando la vista atrás, ya en 2017, el informe TIC360 que publicábamos dedicado a la transformación digital de la universidad, contenía la siguiente gráfica, de nuestro compañero Lluís Alfons Ariño, con la visión global del conjunto de plataformas que deberían cubrir los cinco aspectos claves en una organización digital:

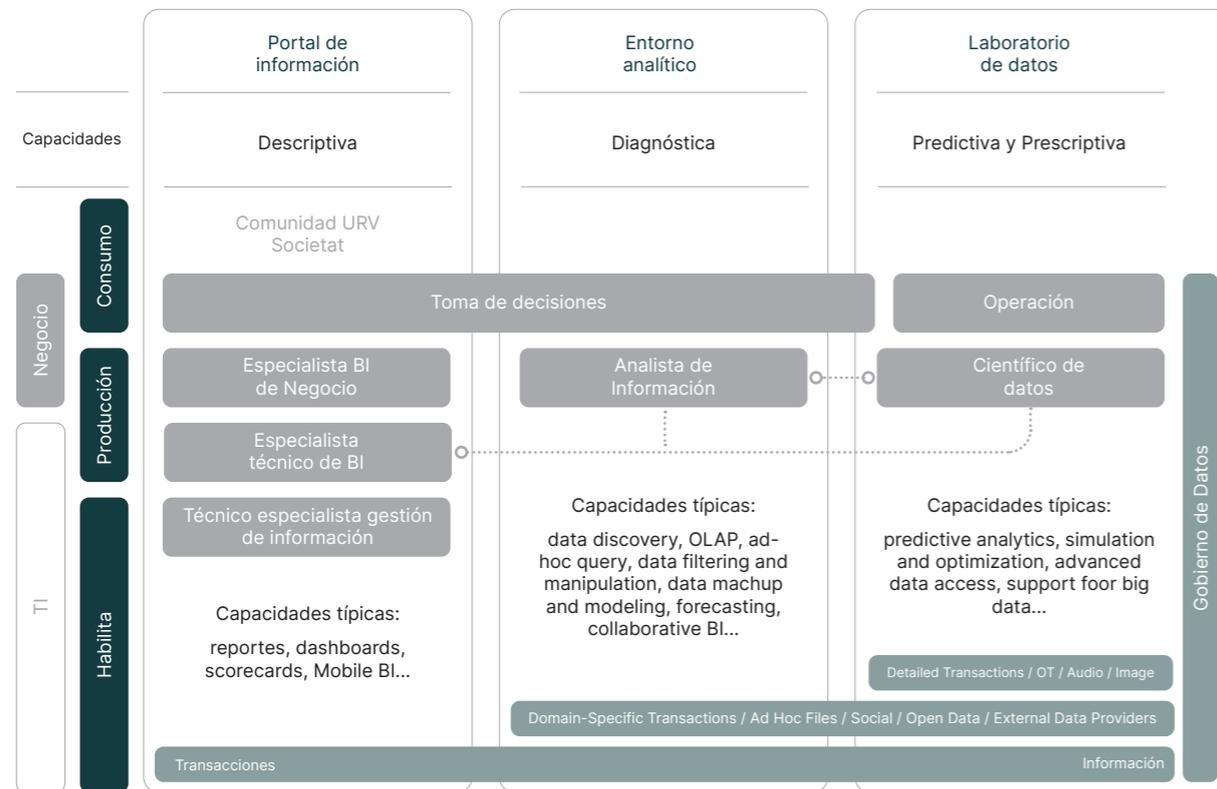
Visión gráfica de los espacios europeos de datos a diciembre de 2022



Observamos ahora, a 6 años vista (una eternidad en la escala de tiempo de las tecnologías actuales), cómo ya situábamos la plataforma de datos y análisis en el centro de ese ecosistema. La plataforma se definía como el conjunto de elementos que han de permitir a la organización la toma informada de decisiones.

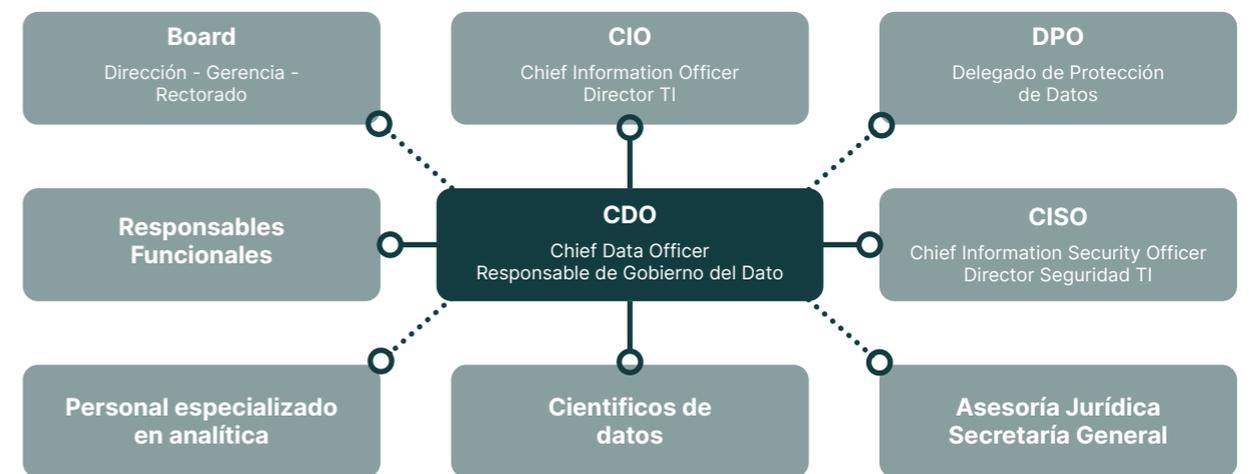
En el informe TIC360 citado se incluye un modelo conceptual basado en el caso de uso de la Universidad Rovira i Virgili (URV) que permite alinear los elementos TI con las capacidades analíticas de la organización.

FIGURA 1.



En el mismo número se incluía un artículo de Santiago Portela denominado «Economía del dato en las universidades» en el que también se situaba al Responsable de Gobierno del Dato como la figura central que interactúa con los diversos órganos de la universidad:

FIGURA 2.



Una referencia más reciente es el caso publicado por el Open Group «Spanish Higher Education Enterprise Architecture Initiative and Capability Map». El documento desarrolla un completo mapa de capacidades necesarias para las Universidades Españolas. La referencia facilita tener un futuro modelo de datos común de las universidades españolas.

Echamos esta breve mirada atrás para reivindicar las continuas llamadas de atención que desde el grupo de directores TI se vienen haciendo acerca de la necesidad de abordar en profundidad este debate en las universidades.

Los datos solo representan propiedades de objetos, hechos o conceptos, y por sí mismos suelen tener poca relevancia, pero cuando los procesamos y los analizamos es cuando estamos en condiciones de extraer información para, finalmente, generar

conocimiento y comprensión. Y este es precisamente el objetivo que se fija la analítica de datos: la extracción de conocimiento a partir de los datos disponibles. Un conocimiento que no se limita a explicar lo que ha ocurrido (análisis descriptivo), sino a predecir comportamientos futuros (análisis predictivo) e incluso recomendar acciones concretas (análisis prescriptivo). Atendiendo a esta definición, se trata de una actividad que los humanos llevamos ejerciendo desde nuestros orígenes, solo que ahora han cambiado algunas cosas.

Aunque el germen de esta disciplina se remonta a los inicios de la estadística, en las últimas décadas, se ha producido un desarrollo muy acelerado facilitado por dos circunstancias:

Por un lado, la gran cantidad de datos disponibles y la posibilidad de su almacenamiento, ya no solo datos estructurados procedentes de las operaciones de las organizaciones, sino también los obtenidos a partir de una gran variedad de medios y formatos generados por nuestra actividad en internet, especialmente de las redes sociales, plataformas de contenido, comercio electrónico y búsqueda de información.

Por otro lado, la enorme capacidad de cómputo y la aparición de nuevas técnicas y herramientas de mano del aprendizaje automático o de la inteligencia artificial, lo que posibilita descubrir patrones y relaciones complejas entre los datos que pasan desapercibidas para los humanos y que nos permiten tomar decisiones informadas basadas en estos hallazgos.

En este contexto, y a partir de la evidencia de la importancia que adquieren los datos, tanto para el desarrollo económico como social, desde la Comisión Europea se trabaja intensamente en establecer las bases necesarias para el desarrollo de una estrategia que nos permita liderar la denominada economía del dato a la vez que se protegen los derechos sobre el control y el uso de nuestros propios datos.

Estos movimientos han tenido un impacto también en las universidades, donde tenemos que provocar el debate sobre el uso de la analítica de datos aplicada a la mejora de la docencia universitaria. Por un lado, se posibilita la detección de patrones de rendimiento académico y, por otro, se facilita la personalización del proceso de aprendizaje. Sin embargo, es necesario ser cautelosos, ya que debemos tener en cuenta los riesgos inherentes a estas herramientas, aún en fase de desarrollo, como son la introducción de sesgos basados en la edad, procedencia, sexo, estatus socioeconómico y otros factores. Existen, además, otras amenazas potenciales, como la posible pérdida de valor de habilidades importantes como la creatividad y la colaboración, en función de los parámetros en que basemos nuestros análisis.

La personalización del aprendizaje también puede llevar a cierta rigidez en la planificación del currículo del estudiante que dificulte el descubrimiento de otros intereses, de forma análoga a como ocurre actualmente con las recomendaciones que nos hacen desde las plataformas de contenido.

La aplicación de la analítica de datos para optimizar procesos de gestión o mejorar la eficiencia energética mediante la monitorización de las infraestructuras presenta menos problemas en términos éticos de privacidad y seguridad que su aplicación a la docencia. En estos casos, la información se obtiene principalmente de sensores y datos administrativos almacenados en los sistemas de información corporativos, lo que reduce las implicaciones negativas de su uso. De esta forma, podemos tomar decisiones más informadas, identificar y resolver problemas, optimizar los procedimientos, ahorrar costes y, en última instancia, mejorar nuestro servicio público de educación superior.

Aunque esta herramienta se está convirtiendo en esencial en la toma de decisiones, hay que advertir de que la tarea no es sencilla en ninguna de sus etapas, y de que cada una de ellas presenta sus propios desafíos. Desde la recopilación de datos has-

ta su divulgación, es necesario abordar cuestiones como la depuración, validación, almacenamiento, tratamiento y explotación, que pueden ser complejas y requerir un enfoque muy cuidadoso, por lo que tendremos que dedicar el tiempo y los recursos necesarios para abordarlas de manera adecuada.

Y una última reflexión sobre algunos problemas, no menores: la privacidad, la protección de los datos personales, la transparencia y la explicabilidad.

Privacidad y protección de datos personales son dos condicionantes muy fuertes que deben ser tenidos en cuenta desde el inicio del diseño, prestando especial atención al origen de los datos, a los consentimientos necesarios, a la obligada confidencialidad y a la divulgación y uso de los resultados por parte de las personas autorizadas. En este aspecto hay que destacar la reciente publicación de la guía Aproximación a los espacios de datos desde la perspectiva del RGPD de la Agencia Española de Protección de Datos, en la que se analiza la creación y el uso de los espacios de datos en relación con la normativa de protección de datos personales.

Transparencia y explicabilidad son requisitos fundamentales para asegurar que los datos utilizados en la toma de decisiones no sean incompletos, sesgados o de procedencia ilegítima. Con el aumento de la cantidad y diversificación de fuentes de datos, esta garantía se convierte en una tarea ciertamente compleja. Además, es crucial comprender cómo estos sistemas generan sus conclusiones, ya que, en caso contrario, se puede erosionar la confianza en los resultados ofrecidos y en las recomendaciones de la analítica. Este desafío puede ser particularmente complejo cuando se procesan cientos o miles de parámetros.

A pesar de todas estas dificultades, concluimos que la analítica de datos tiene un enorme potencial en el ámbito universitario, teniendo siempre presente implementarla con precaución y de manera responsable y ética para evitar riesgos legales y de reputación para nuestras instituciones.

En este número de TIC360 se incluyen artículos que nos ayudarán a comprender mejor la importancia del impacto de la analítica de datos, tanto en las organizaciones en general como en la universidad en particular. Se abordarán cuestiones como la necesidad de crear un mercado único de datos compartido como parte de la estrategia de la Comisión Europea, presentando el concepto de espacio de datos como pieza clave habilitadora de la economía del dato. Se resaltarán la importancia de las certificaciones en el ámbito de los datos detallando los diferentes marcos de certificación relevantes existentes hoy en día en el dominio de los datos. Se expondrá la vinculación de la estrategia de datos y los bancos europeos de datos con diferentes aspectos del Mercado Único Digital, haciendo referencia al proyecto Data Spaces Support Centre (DSSC) y al Digital Space for Skills (DS4Skills). Se destacarán las nuevas organizaciones dirigidas por los datos, organizaciones data-driven, y la importancia de que las universidades se adapten a la era digital convirtiéndose en organizaciones data-driven para poder seguir siendo relevantes en la sociedad. Finalmente, se reflexionará sobre la posición de la Unión Europea en relación con la transformación digital en el ámbito económico y empresarial y la necesidad de una ronda de innovación que transforme profundamente a la Universidad, advirtiendo al mismo tiempo de un escenario de riesgo referido a la falta de cultura de cumplimiento normativo en las universidades españolas en cuanto a la protección de datos personales se refiere.

El buen uso de los datos no nos llevará al paraíso, pero sí puede contribuir a construir una sociedad más sostenible, justa e inclusiva. Por el contrario, su mal uso sí podría acercarnos a un infierno digital. Y como advierte Harari, no se trata de predecir el futuro, sino de plasmar las distintas posibilidades que ofrece.

1. LA OPORTUNIDAD DEL CONCEPTO DE ESPACIO DE DATOS

Carlos Alonso Peña

Director de la División Oficina del Dato
Secretaría de Estado de Digitalización e Inteligencia Artificial

La creación de un mercado único de datos compartidos constituye un elemento clave en la estrategia de la Comisión Europea definida en su comunicación Una estrategia europea de datos¹ publicada en febrero de 2020. En dicha estrategia, la Comisión se marca como objetivo el impulso de la economía del dato de una forma que se ajuste a los valores europeos de autodeterminación en la compartición de datos (soberanía), confidencialidad, transparencia, seguridad y competencia justa. Esto resulta particularmente crítico en un momento en el que se prevé que la economía del dato alcance el 4,9% del PIB europeo en 2025².

1 Comisión Europea - Una Estrategia Europea de Datos

2 Comisión Europea - European Data Market study 2021-2023 (mayo 2021)

3 European Commission. La construcción de una economía de datos europea

4 Una Europa adaptada a la era digital

La construcción de una economía del dato europea³ forma parte del objetivo de creación de un mercado único digital expresado en la estrategia Una Europa adaptada a la era digital⁴. Los espacios de datos comunes europeos son un recurso fundamental en la estrategia europea de datos como habilitadores del impulso de la economía del dato. El reto es fomentar la circulación de los datos en beneficio de todos poniendo en común los datos europeos en sectores clave creando espacios de datos comunes e interoperables.

El dato y su papel esencial en el desarrollo de tecnologías disruptivas como la Inteligencia Artificial supone el factor diferencial de una revolución industrial y tecnológica que nos va a permitir consolidar una economía digital más justa, inclusiva y en línea con los Objetivos de Desarrollo Sostenible y la Agenda 2030. Una auténtica economía del dato con la vocación de nutrir con todas las garantías el desarrollo de dos procesos clave y estratégicos para la reconstrucción de nuestro país: la transformación digital y la transición ecológica.

El Gobierno de España está trabajando intensamente para crear el entorno legal, político y de financiación propicio al despliegue de la economía del dato, a través de las distintas iniciativas que se detallan en la estrategia España Digital 2026 y se despliegan en la Estrategia Nacional de Inteligencia Artificial, en el Plan para la Conectividad y las Infraestructuras Digitales y en la Estrategia para el Impulso de la tecnología 5G. Todas estas prioridades forman parte del Plan de Recuperación, Transformación y Resiliencia que aprovechará la oportunidad de los fondos NextGenEU para su impulso.

A lo largo de las próximas líneas se presenta el concepto de espacio de datos como pieza clave habilitadora de la economía del dato, se avanza en su visión conceptual y se enumeran sus principales características intentando centrar y aterrizar su concepto, todo ello antes de vincular su contenido con el sector público y con el mundo académico, detallando en este caso el estado actual del espacio europeo de ciencia abierta y el espacio europeo de habilidades.

5 Orden ministerial ETD/803/2020

El dato es, sin lugar a duda, el gran protagonista y el activo vertebrador de la transformación digital. La infraestructura esencial de la revolución industrial del siglo XXI que ha irrumpido en nuestras sociedades, acelerando un crecimiento exponencial de la generación y el consumo de datos e información sin precedentes que se augura siga adelante en un ascenso explosivo e imparable. Los datos son el centro de la nueva economía y tienen valor económico dado que permiten que los individuos tomen decisiones con mayor valor y utilidad esperada de la que habrían tomado en ausencia de información.

La economía del dato puede definirse como el conjunto de iniciativas, actividades y proyectos cuyo modelo de negocio se basa en la exploración y explotación de las estructuras de datos existentes (tradicionales y procedentes de nuevas fuentes) con el objetivo de identificar oportunidades de generación de productos y servicios. La economía del dato captura un concepto más amplio que el mercado de datos, ya que considera el valor y la riqueza generados en la economía en su conjunto (no solo entre las empresas) por la explotación de datos.

El espacio de datos es pieza clave para el desarrollo de la economía del dato, al habilitar su acceso, intercambio y reutilización legítima, posicionando al dato como recurso no rival, cuya utilidad crece según su uso se generaliza en un claro ejemplo de efecto red. Así, se podrá avanzar hacia una colectivización del valor generado en los ecosistemas de datos. Ecosistemas que se basan en la comunidad, la transparencia, la innovación y la capacidad de escalar y generar beneficios compartidos. En contraposición a los modelos de plataforma, en los que gran parte del valor se retiene en la intermediación, se apuesta por un modelo de ecosistemas que permita a los participantes mantener su autonomía y al mismo tiempo colaborar en transacciones punto a punto

Un espacio de datos es un ecosistema donde materializar la compartición voluntaria de los datos de sus participantes dentro de un entorno de soberanía, confianza y seguridad, establecido mediante mecanismos integrados de gobernanza, organizativos, normativos y técnicos. Los espacios de datos van más allá del intercambio bilateral de información, constituyendo en su versión más avanzada auténticas redes de compartición donde materializar el valor del dato, y garantizando la equidad de su reparto entre empresas, ciudadanos y organismos públicos.

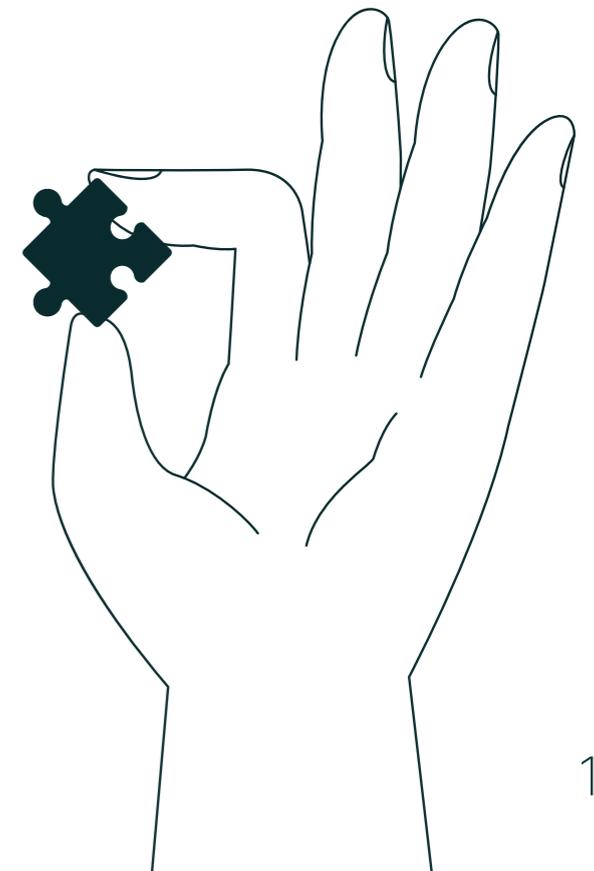
En esta línea, la Oficina del Dato dentro de la Secretaría de Estado de Digitalización e Inteligencia Artificial incluye dentro de sus competencias⁵ «la creación de espacios de compartición de datos entre empresas, ciudadanos y Administraciones Públicas», constituyendo esta línea de acción uno de sus principales cometidos. La Oficina coordina la adaptación de estos nuevos paradigmas de gestión, compartición y explotación de datos, apoyando el desarrollo de sistemas que instancien dichas metodologías dentro de los marcos regulatorios competentes, buscando garantizar la soberanía digital, considerando los distintos modelos y procesos de negocio involucrados.

6 https://ec.europa.eu/info/funding-tenders/find-funding/eu-funding-programmes/digital-europe-programme_en

Aunque entornos de compartición de datos existen desde hace tiempo, la creación de espacios de datos que garanticen los valores y principios de la UE y desarrollados en iniciativas legislativas, constituye no solo un reto tecnológico, sino también de coordinación entre los participantes, de gobernanza, de adopción de estándares y de interoperabilidad.

El reto es de tal magnitud que la Comisión tiene previsto invertir en el 2027 hasta casi 8.000 millones de euros⁶ en el despliegue de la transformación digital europea, en el cual se incluye el fomento de infraestructuras, herramientas, arquitecturas y mecanismos para la compartición de datos. Para que triunfe esta estrategia, es necesario que surja, desde el cumplimiento con los valores europeos, un paradigma de espacio de datos que arraigue en la industria, que se traduzca en un estándar tecnológico de facto, que además avance la concienciación general sobre las posibilidades del dato, y posibilite el retorno económico de las inversiones requeridas.

Las diferentes iniciativas, arquitecturas y modelos de referencia para la creación de espacios de datos en Europa (IDSA, Fiware, Gaia-X y BDVA) se encuentran en proceso de confluencia. En este sentido, el proceso de convergencia tecnológica se encuentra ya en desarrollo tanto por medios públicos (e. g. a través de la iniciativa Simpl, o la financiación del Data Space Support Center) como privadamente (Data Space Business Alliance), y busca hacer confluir los principios y la arquitectura de Gaia-X con el modelo arquitectónico de IDS, y con los bloques constructivos preexistentes de Fiware



7 <https://data.europa.eu/doi/10.2760/400188>

8 Data Sharing Coalition. Data Sharing Canvas

9 International Data Spaces. Enabling Data Economy

10 BDVA TOWARDS A EUROPEAN-GOVERNED DATA SHARING SPACE

11 i4 Trust. Data Spaces for effective and trusted data sharing

12 Gaia-X Vision & strategy

13 European Commission. Staff working document on data spaces

La noción de soberanía es clave dentro del concepto de espacio de datos, entendiéndose como la capacidad de un participante de mantener el control sobre sus propios datos, expresando los términos y condiciones que regirán sus usos permitidos. Así, un proveedor de datos, formalmente incorporado al correspondiente espacio de datos, hará accesibles sus productos de datos y sus condiciones de uso a través del catálogo de datos gestionado directamente por el operador del espacio de datos o con mediación de los oportunos intermediarios de datos. Cuando un partícipe desee acceder a un conjunto de datos, buscará si dicha información se encuentra disponible en el servicio de intermediación de datos, procediendo al estudio de sus condiciones de acceso y uso, así como la idoneidad de su semántica y vocabulario, explicitadas en el oportuno catálogo de datos. Si las características detalladas en el catálogo de datos satisfacen sus expectativas, se iniciará la oportuna negociación, y, llegado el caso, se procederá a establecer el dialogo con el proveedor de datos, generándose una comunicación efectiva entre proveedor y consumidor de datos de acuerdo con las condiciones técnicas explicitadas.

Siguiendo el posicionamiento político analizado en el documento del JRC European Data Spaces - Scientific Insights into Data Sharing and Utilisation at Scale⁷ y considerando la bibliografía al respecto^{8 9 10 11 12 13}, es posible establecer un conjunto de características que idealmente deben cumplir los espacios de datos, agrupadas desde un punto de vista conceptual, operacional o tecnológico.



Principios rectores y legislación

14 DG Connect. Technical guidance report on data security

1. El espacio de datos habilita el acceso e intercambio de datos, no referido al intercambio de su propiedad, promoviendo la **creación continua y sostenible de nuevos productos y servicios** alrededor del dato, completando su cadena de valor, habilitando la competencia justa, reduciendo barreras de entrada al mercado, impulsando la innovación y generando impacto y beneficio en la sociedad
2. El espacio de datos apoya el cumplimiento específico de la **legislación de la Unión Europea**, en particular la protección de datos personales, la legislación de protección del consumidor y las leyes de defensa de la competencia, así como la regulación relativa al tratamiento de datos horizontal y sectorial que le sea de aplicación
3. El espacio de datos fomenta la **confianza de los partícipes**, más allá del cumplimiento normativo y del seguimiento de los valores europeos, generando relaciones de valor añadido entre los partícipes y habilitando una ética de tratamiento del dato alrededor de los principios de administración confiable, privacidad, transparencia y explicabilidad, equidad en el uso de inteligencia artificial y régimen justo y responsable del dato
4. El espacio de datos, como ecosistema de compartición de datos, garantiza la **soberanía de los partícipes** sobre sus datos. El partícipe en el espacio es quién decide con quién intercambia información y en qué condiciones contractuales lo hace, haciendo hincapié en la prevención de usos no autorizados por terceras partes
5. El espacio de datos presta especial atención a todos los aspectos relativos a su **seguridad**¹⁴, incidiendo en su seguridad desde el diseño. Si su criticidad así lo requiere, podrá ser objeto de la correspondiente certificación. Apoyando esta seguridad, se establecerían sistemas de identificación de participantes, así como la idoneidad de los componentes software empleados, mediante mecanismos de homologación o certificación



Participantes y roles

14 DG Connect. Technical guidance report on data security

6. El espacio de datos está abierto a la participación de todo actor del ecosistema, desde el sector público o privado hasta el tercer sector, comprometido a cumplir con sus líneas de funcionamiento. En particular, un espacio de datos concreto guiará su funcionamiento por un código de gobernanza explícito y público, garantizándose su mayor transparencia, con especial atención a la preservación de la equidad de participantes y su no discriminación, así como su sostenibilidad en el tiempo

7. El espacio de datos permite a los diversos participantes el despliegue de diferentes roles dentro del mismo, bien sea como productores de datos, consumidores de los mismos, proveedor de servicios sobre datos, desarrolladores de componentes u operadores de servicios esenciales.

El espacio de datos facilita el desarrollo del rol de intermediario de datos mediante la prestación de los oportunos servicios de intermediación de datos. En particular, habilitará la creación de la figura del tercero de confianza donde, llegado el caso, poder realizar el tratamiento de la información facilitada por las partes. Los espacios de datos dispondrán un operador del espacio de datos, imparcial y neutral, que gestione los servicios necesarios, considerando la necesidad de que haya en cualquier caso un consorcio de operación responsable en última instancia del servicio.



Gobernanza

15 European Commission. Turning fair into reality

8. El código de gobernanza de un espacio de datos habilitará la consignación de políticas de acceso y uso de la información en los correspondientes acuerdos, pudiendo establecer compensaciones económicas u otro tipo de incentivos, vigilándose su proporcionalidad y la no existencia de costes excesivos de transacción, y primándose su consignación electrónica. En caso de conflicto, se podrán gestionar a través de mecanismos de resolución de conflictos estipulados en el código de gobernanza

9. El espacio de datos facilitará encontrar, acceder y usar los datos, bajo principios FAIR¹⁵ (localizables, accesibles, intercambiables, reusables), describiendo suficientemente los conjuntos de datos implicados y sus restricciones de uso, las estructuras de datos, vocabularios y taxonomías, así como los medios técnicos de acceso. Dicha descripción se realizará, en la medida de lo posible, siguiendo las indicaciones de los organismos de normalización o, en su ausencia, los criterios comunes dictados por las instituciones europeas para la interoperabilidad de espacios de datos

El espacio de datos hará un uso intensivo de las fuentes de datos abiertos aportadas por el sector público, considerando sus condiciones generales de licenciamiento o contractuales específicas de aplicación. En especial, considerará la potencialidad de los conjuntos de datos de alto valor (HVDS). El espacio de datos dispone de mecanismos para asegurar la transparencia sobre el nivel de calidad de los datos aportados por los participantes.

10. El espacio de datos facilita el desarrollo de semánticas y vocabularios de uso común facilitadoras del intercambio de información. El mantenimiento de estas semánticas deberá ser objeto de observancia por el consorcio de operación

11. Los datos recopilados en los distintos espacios de datos, cuando estén disponibles para el público, se presentan en formatos accesibles para las personas con discapacidad en igualdad de condiciones con los demás ciudadanos

12. El clausulado contractual de un caso de uso del espacio de datos podrá fijar la disponibilidad del producto final resultante de la compartición a parte o a la totalidad de partícipes, o su publicación bajo licencia de datos abiertos. La información así generada debe ser ofrecida en las mejores condiciones para su identificación, acceso, descarga y análisis posterior

13. El funcionamiento del espacio de datos en su conjunto, así como el uso dado a los datos por los diferentes participantes, podrá ser objeto de auditoría externa, en particular todo lo relativo a la verificación del cumplimiento de las políticas de incorporación y los acuerdos de acceso y uso de los datos. Para ello, y por otros motivos, se deberían generar oportunas trazas de funcionamiento, garantizándose el acceso de cada partícipe a su parte correspondiente de información

14. El espacio de datos asegurará la interoperabilidad y sinergias con otros espacios de datos sectoriales nacionales o europeos





Tecnología

15. El espacio de datos pretende ser independiente de la solución tecnológica subyacente, permitiéndose de hecho su portabilidad y despliegue en diferentes infraestructuras físicas, incentivándose de esta forma la generación de un ecosistema innovador alrededor de soluciones y servicios

16. Los componentes *software* presentes en los espacios de datos se crearán bajo esquemas de licenciamiento que habiliten su conocimiento y reutilización. Se obtendrán, siempre que sea posible, desde repositorios de código de fuentes abiertas y cualquier adaptación realizada sobre los mismos deberá ser accesible a la comunidad de desarrollo

17. El espacio de datos hará uso, allí donde sea necesario, de tecnologías innovadoras, pudiéndose reseñar las siguientes situaciones de utilización:

- Se facilitará el despliegue de servicios y herramientas de analítica avanzada (descriptiva, predictiva, prescriptiva) con las que obtener valor adicional de los datos intercambiados
- Se potenciará el uso intensivo de tecnologías de privacidad (PET) para garantizar el adecuado tratamiento de información protegida o confidencial sometida a propiedad intelectual o secreto comercial
- Se potenciará el despliegue y uso intensivo de modelos innovadores de aprendizaje federado, facilitando la ejecución de algoritmos de extracción de conocimiento de uno o varios participantes, sin precisar el acceso directo a los datos origen del citado conocimiento
- Se podrán utilizar las tecnologías de registro distribuidas para garantizar tanto la propiedad de los datos en sí, como de sus derechos de acceso, buscando evitar la dilución de su valor
- La negociación del contrato de adhesión al espacio, así como las condiciones de uso de los datos podrá realizarse mayormente en línea, tomando partido por tecnologías relativas a smart contracts
- Se aprovecharán las tecnologías **tipo digital wallet** (PIMS) en donde el ciudadano identificado digitalmente decide qué uso se hace de la información facilitada por él

Las características enunciadas no hacen sino detallar la definición inicial de espacio de datos como entorno de soberanía, confianza y seguridad donde habilitar la compartición de datos, más allá de la solución tecnológica, abarcando aspectos de gobernanza, organizativos, normativos y técnicos.

El sector público ante los espacios de datos

Los datos, entendidos como bien público, son una pieza clave en el proceso de transformación digital de las administraciones públicas, redefiniendo su relación con la ciudadanía y los diferentes sectores productivos, buscando siempre potenciar el bien común de la sociedad y una economía justa e inclusiva.

El dato es un bien público a conservar y tratar, en aras de una aplicación de servicios y políticas públicas de calidad. Un uso intensivo de datos permite impulsar la innovación en la actuación del sector público, facilitando el contraste de ideas, promoviendo la creatividad y el máximo aprovechamiento de los recursos en el marco general de una gestión pública moderna, participativa, abierta y útil para solucionar o mejorar los problemas y retos sociales.

El objetivo es, asumiendo un enfoque mucho más interdisciplinario e interdepartamental, y sacando partido de las últimas tecnologías, proyectar a mayor escala las metodologías, especificaciones y prácticas actuales relativas al tratamiento de información; lograr un intercambio de datos fluido y continuo entre administraciones, sectores industriales y ciudadanos, que genere ventajas y oportunidades a los diferentes actores implicados, y teniendo siempre presentes las necesarias consideraciones de privacidad y seguridad.

Este fortalecimiento de la colaboración administrativa se materializará en el espacio de datos del sector público, potenciando el valor del dato en la elaboración de programas, políticas y servicios públicos centrados en el ciudadano, reduciendo la carga burocrática de los procesos administrativos soportada por los operadores económicos y los ciudadanos.

Los datos son una oportunidad para facilitar la concepción y aplicación de políticas públicas basadas en evidencias; adoptar decisiones informadas tomando como base datos más precisos y actualizados; prestar más y mejores servicios con mayor orientación al ciudadano, garantizando su eficacia y permitiendo medir su efectividad; ayudar a crear valor o facilitar la investigación ofreciendo la información para su reutilización, y favorecer la transparencia en la gestión pública.

La información atesorada por las administraciones debe ser accesible para fines de investigación social o análisis de políticas públicas. Los investigadores, debidamente autorizados por la autoridad competente, deben poder acceder a la información de forma ágil, más allá de las salas físicas y de la información presente en una única organización, garantizándose en todo caso la privacidad de la información tratada y la imposibilidad de su reconstrucción.

Las administraciones deben convertirse en parte fundamental de los espacios de datos sectoriales. Es necesario establecer nuevas asociaciones basadas en datos entre las administraciones y la industria, fomentando la cultura de datos abiertos con los que la industria pueda desarrollar nuevos modelos de negocio. Los datos abiertos son pieza clave de los espacios de datos, surgiendo el rol del intermediario de datos oferente de servicios de valor añadido sobre la información cruda facilitada por las administraciones. La colaboración entre administraciones y con el sector privado y la sociedad civil es imprescindible para completar la cadena de valor de los datos, alentando el dinamismo de la iniciativa privada y de la sociedad civil a la hora de la creación de nuevos productos y servicios de valor añadido basados en los datos.

El mundo universitario y académico ante los espacios de datos

El mundo académico debe ver a los ciudadanos, la sociedad civil y las empresas como socios necesarios en su camino de transformación digital, transformación que inevitablemente pasa por su trayecto hacia una organización orientada al dato. El mundo universitario y académico debe estar adecuadamente concienciado acerca de la centralidad del dato y los retos y oportunidades que plantea, alcanzando a ser referente en el uso intensivo del dato. El objetivo es hacer accesible los datos e información necesaria, tanto en el proceso de toma de decisiones interna como para fines de investigación científica en el ámbito social.

Bajo el marco de la estrategia europea de datos, en la cual se están creando y desarrollando espacios de datos en campos estratégicos como salud, agricultura, manufactura, energía, movilidad, finanzas o administraciones públicas, también hay cabida para los espacios de datos relacionados con el mundo académico, como son el espacio de datos de habilidades o el de ciencia abierta. El objetivo final es alcanzar la interoperabilidad de estos espacios de datos sectoriales, interoperabilidad habilitadora y potenciadora de un auténtico mercado único de datos.

Ambos espacios de datos, habilidades y ciencia abierta se encuentran, al igual que el resto de los espacios de datos europeos, en una situación preliminar de inicio de trabajos, lo que no es óbice para que su estado deba ser objeto de seguimiento continuo por los diferentes actores interesados en general, y por el mundo académico en particular. A continuación, se expone la situación de ambos espacios de datos.

Espacio Europeo de Ciencia Abierta (EOSC)

15 Ley 17/2022, del 5 de septiembre de 2022, de la Ciencia, la Tecnología y la Innovación. (Art. 37: Ciencia Abierta).

17 Proyecto de Ley Orgánica del sistema universitario 121/000111

La ciencia y el conocimiento en abierto aumentan la transparencia y fomentan la participación, la cooperación, la rendición de cuentas, la capacidad de reutilización del trabajo investigador y la reproducibilidad de resultados, mejorando así la calidad y fiabilidad de la investigación. La recientemente aprobada la Ley 17/2022, de 5 de septiembre¹⁵ pone en valor la ciencia como un bien común y defiende el acceso abierto a los resultados de la investigación financiados con fondos públicos. Su artículo 37 «Ciencia abierta» impulsa el desarrollo de infraestructuras y plataformas abiertas; obliga al depósito en abierto de publicaciones, datos, códigos y métodos en repositorios y fomenta la participación abierta de la sociedad civil en los procesos científicos. Por su parte, el proyecto de Ley Orgánica del Sistema Universitario LOSU¹⁷, contiene en su artículo 12 «Fomento de la Ciencia abierta y ciudadana» el mandato de depósito en abierto a los resultados de investigación en el ámbito universitario y sitúa a los repositorios institucionales como elementos centrales para su cumplimiento.

18 Portal EOSC

19 Consulta preliminar de mercado EOSC

El Espacio de datos de ciencia, investigación e innovación será pieza clave para el desarrollo de estas posibilidades junto con el desarrollo de repositorios institucionales de acceso abierto a publicaciones científicas y a datos de investigación. El espacio europeo de ciencia abierta, más conocido por su acrónimo en inglés EOSC (Open European Science Cloud), es reconocido por el Consejo de la Unión Europea entre las 20 acciones de la agenda política 2022-2024 del Espacio Europeo de Investigación (ERA, por sus siglas en inglés), con el objetivo específico de profundizar las prácticas de ciencia abierta en Europa.

El objetivo es lograr una federación de infraestructuras que brinde acceso continuo a los diferentes objetos de investigación y servicios de valor agregado durante el ciclo de datos de toda investigación, desde el descubrimiento y la extracción de conocimiento hasta el almacenamiento, la gestión, el análisis y la reutilización a través de fronteras y disciplinas científicas. En última instancia, el EOSC tiene como objetivo desarrollar una «Red de datos y servicios FAIR» para la ciencia en Europa sobre la cual construir una amplia gama de servicios de valor agregado. Estos van desde la visualización y el análisis hasta la preservación de la información a largo plazo o el seguimiento de la adopción de prácticas de ciencia abierta.

Actualmente, existe una puerta de entrada a la información y a los recursos del EOSC, denominada Portal EOSC¹⁸, puerta de entrada a la información y los recursos en EOSC, que proporciona actualizaciones sobre su gobierno y actores, los proyectos que contribuyen a su realización, las oportunidades de financiación, políticas europeas y nacionales relevantes, documentos importantes y desarrollos recientes. El catálogo y marketplace de EOSC actúa como punto de entrada a la multitud de servicios y recursos disponibles para investigadores.

Tras una fase preliminar de implementación (2018-2020) basada en subvenciones de la UE (programa Horizonte 2020), el EOSC actualmente se encuentra en una segunda fase de desarrollo y maduración como Asociación Europea co-programada dentro de la Agenda de Investigación e Innovación Estratégica (SRIA). Se prevé una inversión por parte de los socios de la UE de unos 1.000 millones de euros, entre contribuciones financieras y en especie, a repartir durante el periodo 2021-2027 con el objetivo de mejorar el almacenamiento, el intercambio y la reutilización de datos de investigación a través de las distintas fronteras y disciplinas científicas.

Para capitalizar el esfuerzo de investigación e innovación, la Comisión Europea adquirirá la infraestructura operativa de la EOSC como parte de las disposiciones del Programa de trabajo 2021-2022 de Horizon Europe, Research Infrastructures. Para ello, se ha realizado una consulta preliminar de mercado¹⁹ entre mayo y agosto de 2022 con el objetivo de identificar diferentes enfoques de contratación, los desafíos técnicos y las posibles soluciones. En base a la información obtenida en la citada consulta, la Comisión decidirá si precisa consultas adicionales o si ya dispone de información suficiente para seguir adelante con la contratación de la citada infraestructura. La contratación se prevé lanzar antes de finales de 2022, siendo un proyecto cuya evolución interesa seguir de cerca.

El Espacio de Datos de Habilidades (Common European Skills Data Space) pretende reducir la brecha existente en torno a las habilidades adquiridas en los sistemas de educación y formación y las necesidades reales del mercado laboral, aumentando la productividad y competitividad de los participantes. Hoy en día, la información sobre habilidades, perfiles educativos, ofertas de trabajo y ofertas de capacitación está dispersa y almacenada en silos y distribuida entre una amplia gama de partes interesadas públicas y privadas, dificultando desarrollar su potencial. El Espacio Europeo de Datos de Habilidades quiere permitir que todas las partes interesadas se unan a una red común de intercambio de datos para superar los silos y crear aplicaciones y usos innovadores de los datos, aprovechando oportunidades a nivel individual, organizacional y social.

Así, el espacio proporcionará un fácil acceso transfronterizo a un conjunto de datos clave, desarrollando servicios de datos coincidentes con los valores en un entorno ético, seguro, y soberano, sentando las bases para las grandes transformaciones de los sistemas educativos, apoyando el desarrollo de políticas públicas basadas en datos, su reutilización en aplicaciones innovadoras o la actualización de programas de educación y formación incluyendo las necesidades propias de las habilidades emergentes.

Los fundamentos del espacio de datos (acción preparatoria) se están desarrollando a través del Programa Europa Digital (DIGITAL) de la Comisión Europea, buscando involucrar a las diversas partes interesadas tales como organizaciones europeas, administraciones públicas y organismos gubernamentales, actores privados y públicos, proveedores de educación y proveedores de formación. El adjudicatario del programa, con un presupuesto de un millón de euros, es el proyecto Data Space for Skills²⁰ (#DS4Skills), consorcio de 14 socios liderado por DIGITALEUROPE, abarcando actores clave de la industria, la educación y el ecosistema de datos.

Inicialmente el programa identificará las fuentes de datos relevantes para el espacio de datos de habilidades mediante el mapeo y el análisis de las plataformas y proyectos existentes para recopilar, almacenar y compartir datos de habilidades del sector público y privado. A continuación, el proyecto desarrollará un modelo para la implementación de este espacio de datos en el futuro, con varios enfoques y opciones conceptuales, así como casos de uso específicos. Todas estas actividades se prevén realizar con la oportuna involucración y consenso con las diversas partes interesadas, coordinadamente con el desarrollo del resto de espacios de datos sectoriales.

El concepto de espacio de datos se va a desplegar con todo su potencial a lo largo de los próximos años, dado que son la apuesta clave de la Unión Europea para impulsar la economía del dato y preservar su soberanía digital. El mundo académico y universitario debe tener presente el dato como base de sus actividades, y en concreto debe estar vigilante del concepto de espacio de datos en general y de los espacios de datos de habilidades y ciencia abierta en particular.

Las diferentes iniciativas, arquitecturas y modelos de referencia para la creación de espacios de datos en Europa se encuentran en proceso de confluencia. De la rapidez y agilidad de las organizaciones implicadas en entregar un producto maduro que cumpla las expectativas, dependerá en gran medida el ritmo de despliegue y la interoperabilidad de los diferentes espacios de datos a constituir y por tanto del despliegue del mercado de datos en sí.

Aquí, el mundo de la investigación académica alrededor de las tecnologías fundacionales de los espacios de datos puede jugar un papel importante, si no en la definición de las arquitecturas y bloques constructivos, sí en la realización de los desarrollos, parametrizaciones y ajustes necesarios para su despliegue particularizado, sin olvidar el terreno de juego que se abre en la definición de las semánticas y ontologías, así como normativas y estándares de calidad, precisas para alcanzar su ansiada interoperabilidad.

Más allá de la tecnología, los espacios de datos deben entenderse como el lugar donde crear nuevas oportunidades de servicio público o de negocio, generadoras de sinergias derivadas de la compartición de datos más allá de la suma de los datos individuales. El problema a resolver es un problema de política pública o de negocio, y la solución nunca debe ser exclusivamente tecnológica, deben buscarse soluciones simples con facilidad de adopción y uso, de despliegue robusto, ágil y sencillo que no generen barreras de entrada a pequeños participantes.

El reto de desplegar los espacios europeos de datos interoperables no es sencillo. Son muchos los intereses que están en juego, siendo fundamental su desarrollo para, aprovechando el momento tecnológico, económico, político y social, conseguir una economía orientado al dato, sostenible, generador de valor social e inclusivo. En esta transformación es donde el mundo universitario y académico debe jugar un papel fundamental.

2. EUROPEAN SKILLS DATA SPACE

Lluís Alfons Ariño Martin

Universitat Rovira i Virgili
Miembro del Advisory Board de DS4Skills



Contexto. Digital Single Market

21 <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52015DC0192>

22 https://commission.europa.eu/publications/european-commission-digital-strategy_en

23 https://reform-support.ec.europa.eu/what-we-do/digital-transition_en

24 <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>

25 <https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces>

26 <https://digital-strategy.ec.europa.eu/en/policies/data-act>

27 https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en

28 <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>

29 <https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act>

30 <https://www.europarl.europa.eu/>

31 https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-chips-act_en

32 https://commission.europa.eu/publications/interoperable-europe-act-proposal_en

El Mercado Único Digital (Digital Single Market²¹) designa la estrategia de la Comisión Europea para el mejor acceso posible al mundo en línea para particulares y empresas.

El despliegue de la estrategia se plasma en diversas estrategias y acciones. Así pues, la consecución del Mercado Único Digital puede concebirse como un poliedro con diferentes dimensiones, en el que diferentes estrategias y acciones ayudan a la consecución de subobjetivos que van completando los hitos necesarios para hacer del Mercado Único Digital una realidad en 2050.

En este sentido, podemos vincular/llegar a la estrategia de datos y a los bancos europeos de datos desde diferentes ángulos del poliedro del Mercado Único Digital.

1. Partiendo del pilar de Economía y Sociedad del Mercado Único Digital, este establece un enlace directo con la Economía del Dato, que a su vez establece la Estrategia de Datos
2. Partiendo de la Estrategia Digital²², tenemos también un vínculo directo con la Estrategia de Datos
3. Partiendo de la Estrategia Digital, podemos llegar a la Transición Digital²³, que tiene como objetivos clave la previsibilidad empresarial para planificar la inversión y la necesidad de que las personas sepan que pueden confiar en la tecnología que tienen en sus manos. Para su consecución está definiendo un conjunto de regulación a tener en cuenta también en la Estrategia de Datos²⁴ y los Bancos Europeos de Datos²⁵ (a fecha de diciembre de 2022 la Data²⁶, Digital Market²⁷, Digital Service²⁸, Data Ciber Resilient²⁹, Artificial intelligence³⁰, Europe's Chip³¹, Interoperability³²).

Por una u otra vía se llega a la Estrategia de Datos, en la que se definen explícitamente los Espacios Europeos de Datos. Hasta el momento, a fecha de diciembre de 2022, se han identificado 17 espacios europeos de datos:

FIGURA 3. Visión gráfica de los espacios europeos de datos a diciembre de 2022



Fuente: DG-CONNECT-G.2.

Cabe remarcar que el trabajo realizado por el entorno educativo, en nuestra vertiente de investigación, se ha tomado como una referencia para el resto de los espacios europeos de datos.

Respecto al European Skills Data Space, el Espacio Europeo de Habilidades, es importante subrayar su transversalidad con el resto de los espacios sectoriales, dada la naturaleza del mismo.

Espacio Europeo de Datos de Habilidades

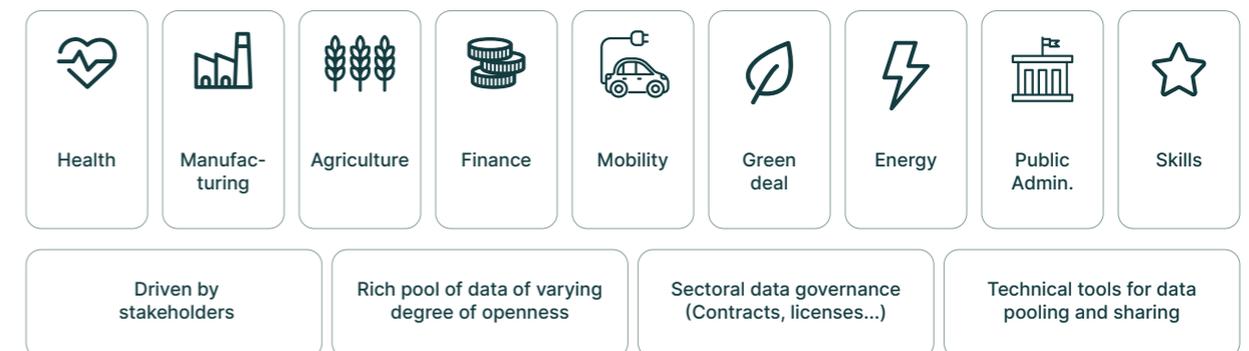
33 <https://dssc.eu>

El desarrollo de los espacios europeos de datos tendrá un germen de coordinación común, desarrollado por el llamado Data Spaces Support Centre³³ (DSSC). Este proyecto tiene como principales objetivos la coordinación del desarrollo de los diferentes espacios europeos de datos, así como asegurar los aspectos necesarios para conseguir unos estándares comunes y la interoperabilidad.

Este centro es un puente entre el desarrollo de los diferentes espacios europeos de datos (que se desarrollan por stakeholders específicos, con diferentes gobernanzas por cada ámbito sectorial, y con diferentes herramientas para llenar y compartir datos), y la infraestructura técnica para los espacios de datos.

FIGURA 4. Visión gráfica de la vinculación/rol del DSSC.

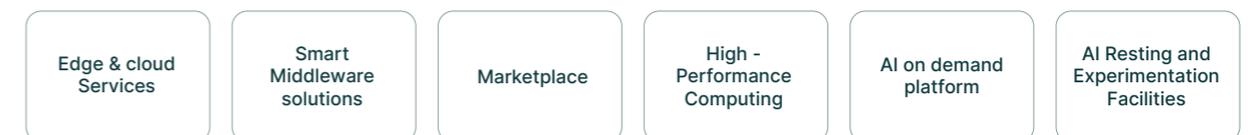
► High Value Datasets from public sector



► Data Spaces Support Centre



► High Value Datasets from public sector



Fuente: DSSC

34 <https://www.skillsdataspace.eu>

Entre otras iniciativas, para facilitar y asegurar la alineación de los espacios de datos, el DSCC facilita herramientas como, por ejemplo, el llamado Starter kit for the data space designers, un toolkit que tiene como objetivos:

1. Identificación de requisitos comunes para la infraestructura de datos en todos los espacios de datos sectoriales
2. Componentes comunes esenciales para la creación de espacios de datos sectoriales
3. Normas comunes, tanto específicas de dominio como transversales
4. Cajas de herramientas comunes que podrían usarse en espacios de datos
5. Arquitectura y marcos técnicos de gobierno de datos
6. Posibles sinergias entre espacios de datos y coordinación de intercambios transversales relacionados entre espacios de datos
7. Promueve modelos de gobierno de datos, modelos de negocio y estrategias para ejecutar espacios de datos
8. Identificar problemas legales y otras barreras relevantes para el mercado
9. Principios rectores de diseño para la creación de espacios de datos

Uno de los proyectos en coordinación por el DSCC es el llamado Digital Space for Skills (DS4Skills³⁴), que sentará las bases del Espacio Europeo de Habilidades.

European Skills Data Space

35 <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018D0646>

36 <https://europa.eu/europass/en/stakeholders/european-digital-credentials>

37 <https://europa.eu/europass/en/node/2128>

38 <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>

Explicado el contexto del Espacio Europeo de Habilidades y la estrategia de datos, veamos específicamente algunos vínculos de esta última con el sector educativo.

Por un lado, en la propia Estrategia de Datos se hace referencia a la Europass Decision³⁵, vinculando las estrategias nacionales de credenciales digitales y referenciando a la iniciativa European Digital Credentials for Learning³⁶ que, a su vez, nos lleva al modelo de datos European Learning Model³⁷. ¿Por qué es esto relevante? Porque el European Learning Model ha de permitir asentar una ontología común sobre la que construir el Espacio Europeo de Habilidades (European Skills Data Space).

Por otro lado, la Estrategia de Datos también referencia explícitamente al llamado EU Digital Credentials Action Plan³⁸ y, obviamente, al European Skills Data Space.

El desarrollo de este espacio se plantea, como mínimo, en dos fases diferenciadas. Una primera fase en la que se realizarán trabajos preliminares para la concepción y definición del espacio europeo de habilidades y una segunda fase en la que, partiendo de los resultados de la primera, se implementará el espacio en sí.

Fase 1 Proyecto DS4Skills

Periodo	Oct'22 to Oct'23 (sujeto a extensión de 2 meses en el momento de redacción del artículo)
Objetivo	Trabajos preliminares para la construcción del Espacio Europeo de Habilidades
Acciones y objetivos previstos	<p>1. Identificación y construcción de un inventario de plataformas existentes. Para ello, se han fijado los siguientes objetivos:</p> <ul style="list-style-type: none">• Abordar la falta de visión general y la fragmentación de las plataformas existentes, debido a la participación de los muchos actores privados y públicos• Garantizar el apoyo de las principales partes interesadas a nivel europeo para que las plataformas tengan en cuenta el diseño del espacio de datos para las competencias• Desarrollar un inventario interactivo en línea que reúna la amplia gama de plataformas existentes relacionadas con los espacios de datos para las habilidades <p>2. Desarrollo del plan para la configuración y el futuro despliegue del espacio de datos de habilidades. Para ello, se han fijado los siguientes objetivos:</p> <ul style="list-style-type: none">• Presentar el diseño y el enfoque conceptual y el plan sobre todos los temas clave (técnicos, gobernanza, modelos de negocio, aplicaciones) que pueden ser la base para el trabajo futuro sobre el espacio de datos para habilidades• Visualizar aplicaciones innovadoras habilitadas por el espacio de datos y que respondan a las necesidades de las partes interesadas, así como a los objetivos políticos, con ejemplos clave que involucren a las partes interesadas identificadas• Validar el enfoque preferido a través del diálogo entre múltiples partes interesadas <p>3. Colaboración con el DSSC</p> <ul style="list-style-type: none">• Alinear el plano para el espacio de datos para habilidades con la arquitectura de referencia y considerar bloques de construcción comunes e interoperabilidad• Alinear los modelos de gobierno de datos, los modelos de negocio y las estrategias para ejecutar espacios de datos con el DSSC• Colaborar con el Centro de Soporte de Espacios de Datos en temas transversales comunes: arquitecturas de referencia, bloques de construcción, cajas de herramientas, estándares, protocolos de interoperabilidad

Fase 2 Implementación del Espacio Europeo de Habilidades

Periodo	Octubre'23 en adelante (sujeto a la finalización de DS4Skills)
Objetivo	Implementación del Espacio Europeo de Habilidades
Acciones y objetivos previstos	A partir de los resultados del proyecto DS4Skills, y bajo coordinación del DSCC, implementar el Espacio Europeo de Habilidades



Visión Europea

39 <https://digital-strategy.ec.europa.eu/en/library/european-declaration-digital-rights-and-principles>

Cabe decir que en el poliedro del Mercado Único Digital también se identifican y establecen un conjunto de principios y derechos que se aplican de manera transversal a cualquier acción. Principios como:

- Europa ahora debe fortalecer su soberanía digital y establecer estándares, en lugar de seguir los de otros, con un claro enfoque en los datos, la tecnología y la infraestructura
- Previsibilidad del negocio para planificar la inversión
- Privacidad aportada por el cumplimiento del GDPR
- Las personas necesitan saber que pueden confiar en la tecnología en sus manos
- Ciudadanos en control total tanto de su identidad como de sus datos

Derechos como:

- Declaración europea sobre los derechos y principios digitales para el Decenio Digital (European Declaration on Digital Rights and Principles for the Digital Decade³⁹)
- GDPR

Dichos principios y derechos conforman unos valores europeos que son un factor diferencial e identificativo de la Unión Europea, que, sin duda, sentarán las bases para el resto de las economías para la protección de los derechos de los ciudadanos y la confianza y creación de oportunidades para las empresas.



Elementos clave del Espacio Europeo de Habilidades

¿Qué contendrá en sí el Espacio Europeo de Habilidades? La respuesta final, obviamente, la ha de aportar el proyecto DS4Skills, pero podemos dar una orientación de algunos posibles contenidos clave a tener en consideración.

Tan importante como los contenidos en sí, será la relación entre objetos, y el establecimiento de una gobernanza efectiva que sustente el modelo.

A modo de ejemplo, podríamos tener los siguientes roles:

- Proveedor de educación y capacitación
- Demandante de empleo
- Ciudadano
- Estudiante
- Empleador

Que podrían aportar al espacio de datos las siguientes contribuciones:

Proveedor de educación y capacitación:

- Detalles de registro
- Detalles personales
- Detalles de la organización
- Detalles del curso
- Detalles de la credencial
- Detalles de habilidades (vocacionales)

Demandante de empleo

- Detalles de registro
- Detalles personales
- Historia de la Educación
- Historial de empleo
- Intereses laborales
- Intereses educativos
- Detalles de habilidades

Ciudadano

- Detalles de registro
- Datos personales
- Historia de la Educación
- Historial de empleo
- Intereses laborales
- Intereses educativos
- Detalles de habilidades

Estudiante

- Detalles de registro
- Detalles personales
- Historia de la Educación
- Historial de empleo
- Intereses laborales
- Intereses educativos
- Detalles de habilidades

Empleador

- Detalles de registro
- Detalles personales
- Detalles de la organización
- Puestos vacantes
- Oportunidades de aprendizaje

Que esperarían de este las siguientes aportaciones:

Proveedor de educación y capacitación:

- Datos personales del solicitante de empleo
- Detalles de educación del solicitante de empleo
- Detalles de habilidades del buscador de empleo

Demandante de empleo

- Detalles personales
- Historia de la Educación
- Historial de empleo
- Oportunidades de empleo
- Oportunidades de educación
- Oportunidades de habilidades

Ciudadano

- Oportunidades de empleo
- Oportunidades de educación
- Oportunidades de habilidades

Estudiante

- Oportunidades de educación
- Oportunidades de habilidades
- Oportunidades de empleo

Empleador

- Datos personales del solicitante de empleo
- Detalle de educación del solicitante de empleo
- Detalles de habilidades del buscador de empleo

40 <https://github.com/emrex-eu/elmo-schemas>

41 <https://emrex.eu>

El año 2023 (con extensión de acciones hasta inicios de 2024) será un año especial ya que se celebra el Year of Skills. Este será, sin duda, un acelerador en este ámbito. Además, en la primavera de 2023 se ha liberado la versión 3.1 del European Learning Model (ELM) que, entre otras innovaciones, pasará a ser JSON nativo. Pero quizás lo más relevante del nuevo ELM será su capacidad de otro modelo de datos ampliamente utilizado en educación, sobre todo en el norte de Europa, como es ELMO⁴⁰. De esta manera, se consolida un paso más para que ELM sea de facto el modelo europeo único de datos para cualquier tipo de aprendizaje (formal, no formal, informal), asentando así una de las bases clave para el European Skills Data Space. Hay un punto importante para que esto pueda ser así, y es que se aplique una gobernanza sobre ELM más allá del actual ámbito de Europass, en la que estén representados los estados miembros, DG-EAC (como el punto de contacto pertinente de educación Primaria, Secundaria y Terciaria), y representantes de EMREX⁴¹ (entre otros actores relevantes). De no ser así, ELMO podría ser la opción más viable.



3. LA UNIVERSIDAD ESPAÑOLA EN EL CONTEXTO DE LOS ESPACIOS EUROPEOS DE DATOS

Ricard Martínez Martínez

Universitat de València
Director de la Cátedra de privacidad y Transformación Digital Microsoft-Universitat de València



La Unión Europea viene impulsando desde principios de siglo la transformación digital de la economía, de la administración y de la sociedad en su conjunto. Si en una primera etapa (eEurope Action Plan) nuestra aspiración fue la de construir una UE conectada en todas sus dimensiones, a medida que la conectividad crece y que los distintos sistemas de información van cumpliendo con los objetivos aparecen nuevos retos estructurales llamados a transformar profundamente nuestra realidad.

En este momento, como han señalado, entre otros los doctores Pedreño y Moreno, en la carrera abierta en el ámbito de la inteligencia artificial y, por extensión, en el mundo digital, parece que la Unión Europea sólo pueda aspirar a alcanzar la medalla de bronce. Basta con verificar cuál es el lugar de origen de las empresas con mayor proyección, capital y volumen de negocio en el mundo para verificar que se encuentran establecidas en Estados Unidos y China.

La Unión Europea se ha erigido en un gigante regulador sin un ecosistema empresarial que lidere el mundo en estos ámbitos. Por ello, debe aprender a dictar normas más eficientes, que impulsen la investigación, la innovación, la transferencia y el emprendimiento y aplicarlas de modo ágil y funcional a la consecución de objetivos relevantes para el bien común. A día de hoy, la realidad no es otra que la ausencia de campeones digitales en la UE con las escasas excepciones de servicios como Spotify. La UE debe recortar distancias y, para ello, abordar el proceso de transformación digital desde la cultura constitucional que le es propia.

La traumática experiencia del siglo XX cimentó en nuestros Estados el convencimiento de la necesidad de acotar desde el derecho la tecnología. Cuando la ciencia se ha empleado para fines militares, o por estados totalitarios, nos ha conducido sistemáticamente al desastre, la destrucción, la pobreza y la pérdida de libertades. Este modelo comporta diferencias de competitividad en relación con la cultura anglosajona, cimentada sobre la idea de una cierta abstención reguladora en la medida en que un sistema de Common Law los conflictos van resolviéndose según se producen. Ello posee desde luego una ventaja innegable, en la medida en la que se difiere a la indemnización por daños cualquier afectación que una tecnología podría producir existen menos cortapisas para un desarrollo sin restricciones. Por otra parte, un enfoque de los derechos fundamentales como límites frente al Estado deja un espacio amplio a la autonomía de la voluntad en las relaciones entre particulares y supone también la existencia de un modelo estatal menos intervencionista desde el punto de vista de la articulación de políticas y garantías de control.

Sin embargo, esta percepción no es del todo cierta. Cuando escándalo tras escándalo, la sociedad norteamericana ha percibido los riesgos del acceso a cierta información, las posibilidades de manipulación social, o el impacto directo en enteras categorías de personas como los menores o los consumidores, ha desarrollado acciones legislativas dirigidas a establecer condiciones de regulación que disciplinan el mercado tecnológico. Pero mientras las regulaciones llegan con cuentagotas, y todavía no existe una ley federal de protección de datos, la monetización de nuestra

vida privada en Internet ha proporcionado a los grandes gigantes norteamericanos el músculo económico suficiente, no solo para erigirse en las primeras compañías del mundo, sino también para hacer frente a las responsabilidades económicas que pudieran derivar de su acción. De hecho, como subraya Mayer-Schöemberger, en la práctica poseen desde el punto de vista financiero y de recursos humanos una mayor capacidad que la propia industria europea para el cumplimiento de las obligaciones normativas que la Unión va estableciendo.

En el caso chino, es evidente que sus reglas del juego son otras. Se trata de un contexto de capitalismo de Estado dirigido por un partido comunista. Se imponen barreras o condiciones de mercado a la entrada de terceros con una estrategia claramente dirigida al aprovechamiento de la innovación de cualquier país del mundo para añadir una capa propia. Y todo ello, se acompaña de una estrategia de inversión masiva del Estado en la promoción de la investigación de la ciencia que convierten a China en el otro gran campeón. Y estas decisiones no solo son de naturaleza económica sino parte de una noción geopolítica muy específica. En este caso, de la mano de un sistema de investigación y transferencia cada vez más consolidado y potente, unido a un conjunto de empresas gigantes con un mercado inmenso, mediante la estrategia de la Franja y la Nueva Ruta de la Seda se busca colonizar otros mercados camino de la construcción de la potencia global dominante a lo largo de este siglo.

La respuesta europea a este escenario pasa por el despliegue de un conjunto de políticas públicas, por un plan Marshall a la europea a través de iniciativas como los fondos Next Generation, que implican un profundo esfuerzo de inversión para digitalizar nuestra sociedad. Funcional a este esfuerzo se encuentra la estrategia de fomento de creación de los espacios europeos de datos, cuyo objetivo primario no es otro que crear un ecosistema de tratamiento de los datos en todos los niveles. Y con ello, potenciar la aparición de un mercado específicamente europeo que tanto en el sector público como el privado sean capaces de abordar todos los segmentos necesarios para la producción, explotación, y monetización de la infraestructura tecnológica europea de la sociedad de la información.



En este contexto, la universidad española constituye un agente fundamental para la transformación digital del país. Por otra parte, tanto en el caso de las instituciones públicas como en el de la universidad privada, debe afrontar su propio proceso de transformación digital. Para ello es relevante tener en cuenta de modo separado la gestión y la investigación.

En primer lugar, los gestores universitarios debemos asumir que, tras los embates de sucesivas crisis económicas, y el envejecimiento sistemático de nuestras plantillas, la Universidad necesita urgentemente de una ronda de innovación que transforme profundamente a nuestras organizaciones. La acusada burocratización de nuestro modelo de gestión es un mantra que, a fuerza de ser repetido, adquiere la categoría casi de verdad indiscutible. Estamos seguros de que pueden aportarse decenas de ejemplos de acciones innovadoras y profundamente transformadoras de distintas universidades. Sin embargo, la percepción general no es otra que sentir que la carga de trabajo de los profesionales universitarios no deja de crecer en tareas repetitivas perfectamente asumibles desde una estrategia de automatización e inteligencia artificial. Y, sin embargo, tareas inanes se hacen descansar sobre la comunidad universitaria en muchos casos a costa del sacrificio de nuestra vida personal, de la conciliación, o del empleo de nuestras capacidades en tareas irrelevantes que perjudican nuestra actividad docente e investigadora.

Podríamos poner más de un ejemplo en el que la transformación digital debería revolucionar nuestro modo de entender las cosas. Y ni siquiera nos referimos con ello a ejemplos disruptivos o de uso masivo de la analítica de datos y de la inteligencia artificial. No es de recibo que seamos capaces de matricular miles de estudiantes en poco menos de cinco días, mientras que la concesión de las becas y ayudas al estudio no sea gestionada y notificada hasta cinco meses después. No existe ni un solo elemento en la concesión de una ayuda al estudio que no pueda ser parametrizado, ni un solo criterio que dependa de subjetividad alguna. Por tanto, nuestra gestión es ineficiente y causa un perjuicio, cuando no angustia, a personas cuya supervivencia económica y cuya capacidad de estudio es altamente dependiente de la ayuda. Debería avergonzarnos que una persona que podría disfrutar de una beca en septiembre, en las 24 horas posteriores a su matrícula, no la reciba hasta un momento indeterminado entre marzo y mayo del año siguiente. En lugar de dedicarse plenamente al estudio, de explotar su talento, nuestro estudiante debe buscar un trabajo y, como consecuencia de ello, el año siguiente no recibirá la misma ayuda al estudio puesto que su nivel de renta ha crecido. Y así, desde una carencia absoluta de imaginación rayana en la estulticia sometemos una vida a un esfuerzo adicional, las más de las veces frustrante y excesivo y afectamos a su rendimiento académico.

Tampoco parece razonable que nuestro profesorado deba dedicar días enteros a formalizar formularios, cuya información se encuentra perfectamente disponible en Internet o en manos de la propia administración universitaria. Resulta sencillamente intolerable que, existiendo procesos e incluso repositorios donde se indexa, por ejemplo, la producción bibliográfica, esta se tenga todavía que mecanografiar en sistemas de información como si viviéramos en la segunda mitad del siglo XX. No me extenderé, porque podrían ponerse muchos más ejemplos de automatización de la analítica del desempeño. Ante este estado de cosas, resulta más acuciante pensar en cómo deberíamos innovar en la gestión de la institución universitaria y en el desarrollo de estrategias ordenadas a transformar nuestra docencia. Mientras andamos discutiendo sobre si la docencia presencial es mejor que la docencia online y sobre

si deben o no deben usarse portátiles u móviles en las aulas, estamos perdiendo la carrera de la innovación, no ya con universidades de otros países, sino con grandes empresas que comienzan a confiar mucho más en sus propios modelos formativos que en un título universitario.

De algún modo, estos ejemplos muestran cómo el reto para la gestión de cualquier institución pública o privada debe pasar de un modelo estático, basado en criterios de planificación a medio y largo plazo, poco coherentes con la velocidad de evolución actual de la tecnología y la sociedad y que a veces dependen de la teoría cuando no de la ocurrencia, a modelos flexibles más adecuados a la realidad líquida que nos envuelve. Necesitamos incrementar la capacidad de modular en tiempo real nuestra estrategia y de entender que no sólo es relevante aprender sino también desaprender. Y sin sacralizar el dato, la realidad es que los modelos de gestión y toma de decisiones asistidos por la analítica de datos adquieren u ofrecen un espacio nunca antes disponible. Ello no implica una renuncia ni al ser humano, ni a nuestra capacidad de tomar de decisiones de naturaleza política o estratégica, pero sí comporta un modo nuevo de hacer las cosas al que debemos prestar atención.

Lo mismo sucede desde el punto de vista de nuestra tarea principal, que es formar personas. En este sentido, la analítica del desempeño y las metodologías de soporte al estudiante pasan ineludiblemente por la transformación digital en nuestra docencia. Es necesario desde un punto de vista proactivo, esto es, a la hora de ofrecer herramientas que completen nuestra misión tradicional, potenciando al máximo las capacidades de nuestros estudiantes. Y también desde un enfoque reactivo, incrementando nuestra capacidad de medir el desempeño del estudiante y de prevenir los supuestos de deshonestidad académica.

Por otro lado, en investigación, la universidad española ocupa, como es bien conocido un papel de particular relevancia. En este ámbito, el crecimiento de la investigación basada en datos se caracteriza por un alto consumo de información que no siempre se encuentra disponible. De hecho, la experiencia demuestra que muchas áreas de investigación más que un entorno de datos abiertos, podemos identificar la presencia de silos de datos. Por ello, la apuesta de la Unión Europea por los principios FAIR resulta altamente significativa. Los datos deben estar disponibles (findable), fácilmente accesibles, deben ser interoperables, y debe fomentarse su reutilización. Funcional a esta filosofía es el impulso para la creación de los distintos espacios de datos. Los datos que genera el sector público, así como aquellos datos generados por el sector privado, cuyo uso se pretende incentivar, constituyen una fuente de riqueza irrenunciable al servicio de la sociedad. De ahí que sea crucial el desarrollo de un mercado europeo de los datos, que permita el acceso a los mismos con fines de investigación, innovación y emprendimiento.

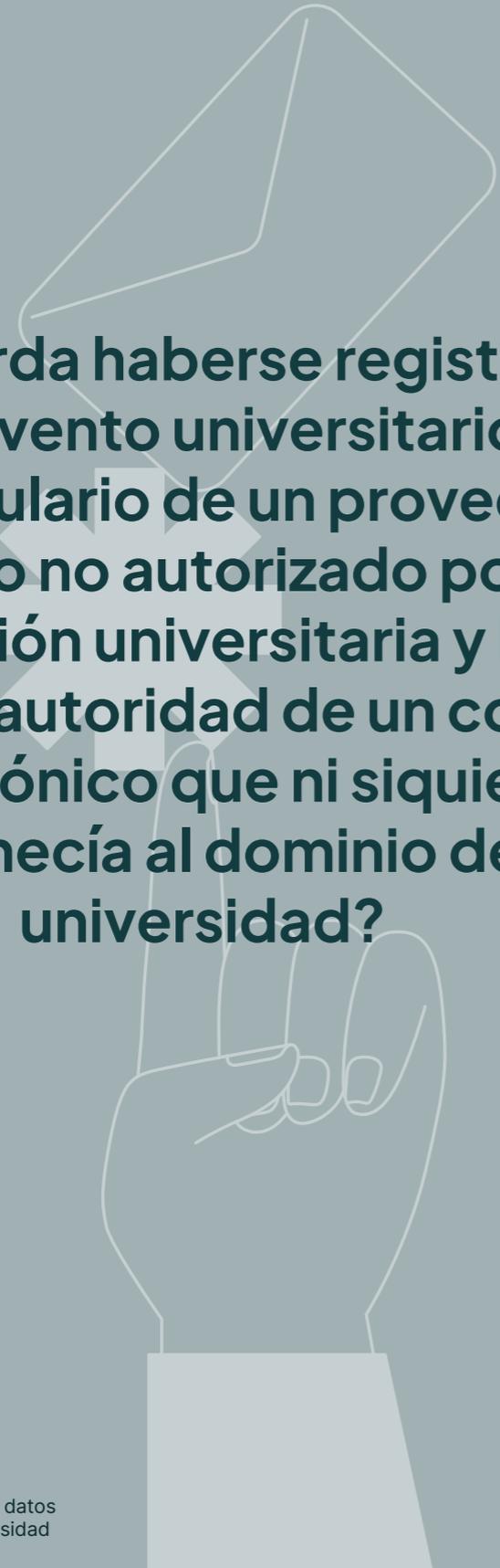
Y esta política de la UE afecta también a la institución universitaria como fuente productora de datos. Nuestras universidades, con contadas excepciones, y en ausencia de una normatividad imperativa que ahora comienza a cambiar, no han sido generalmente entornos facilitadores del acceso a los datos que producen. Usualmente somos generadores de datos de la mano de las obligaciones de compartición en espacios ministeriales o en repositorios de la Unión Europea impuestas por las subvenciones públicas. De hecho, además de con motivo de la investigación, el ecosistema universitario produce conjuntos de datos con motivo de su gestión que podrían producir cierto valor en cuanto a su reutilización. Es evidente que las universidades

generan en su entorno inmediato un impacto económico y transformador. Del mismo modo, la propia actividad de la comunidad universitaria y la gestión de sus intereses define volúmenes de datos significativos para la definición de futuras políticas públicas. Así, no puede entenderse la ordenación de la movilidad en una ciudad si no se tiene en cuenta que, en ocasiones, un campus universitario mueve en horarios muy concretos volúmenes de población de 30.000 o 40.000 personas. Del mismo modo, deberíamos usar nuestros datos académicos para promover un ecosistema favorable a la mejora de capacidades ofreciendo una orientación académica o estratégica adecuada e incluso un soporte vocacional. Todas y cada una de las políticas públicas asociadas a estos ejemplos admiten, al menos conceptualmente, la posibilidad de generarse afinando con mayor precisión nuestro análisis para la toma posterior de decisiones, con el soporte de repositorios de datos propios o de terceros, en combinación con datos públicos o privados.

¿Un escenario de riesgo?

Disculpe lector las consideraciones anteriores, seguramente más propias de un experto en computación, en ciencia política o en gestión de la administración pública. En la práctica no tienen más trascendencia que transmitir aquello que el sentido común de un gestor universitario pueda percibir. Sin embargo, resultaban relevantes para definir escenarios de riesgo. Porque este estado de cosas genera culturas corporativas reacias al cambio y, en lo que a este trabajo interesa, define resistencia a la inversión en perfiles altamente especializados para proporcionar un soporte adecuado a la transformación digital.

La institución universitaria, al igual que el conjunto de la sociedad, ha venido asistiendo desde los años 90 del pasado siglo a las sucesivas leyes orgánicas en materia de protección de datos en un proceso que culmina con el Reglamento General de Protección de Datos. Desde un añejo conocimiento de la gestión de los datos de carácter personal en el ámbito de las administraciones públicas, del marco sancionador vigente en España, y de la experiencia del y en el propio regulador, para nuestra desgracia, la normativa de protección de datos no ha permeado con la suficiente intensidad. La razón última, aparte de una cultura jurídica nacional poco proclive a la idea de la responsabilidad proactiva deriva, sin duda, de la ausencia de incentivo alguno al cumplimiento normativo. En efecto, ley orgánica tras ley orgánica, la interpretación de la Agencia Española de Protección de Datos consolidó, y petrificó legislativamente, un régimen sancionador que no aplicaba a las administraciones públicas. Ello, en un contexto universitario predominantemente público, opera como un elemento que desincentiva por completo el cumplimiento de la norma en la materia.



¿Recuerda haberse registrado en un evento universitario en un formulario de un proveedor privado no autorizado por la institución universitaria y bajo la única autoridad de un correo electrónico que ni siquiera pertenecía al dominio de la universidad?

Y así, al menos hasta la llegada del RGPD, el cumplimiento normativo en este ámbito ha sido puramente epidérmico, reactivo y acompañado de la ignorancia respecto de los procedimientos y el menosprecio de aquellos profesionales que desde la seguridad o la privacidad trataban de poner orden en el caos. La Universidad, desde la autonomía que la caracteriza, no sólo a la institución, sino en muchas ocasiones a las unidades de gestión, centros y departamentos, ha sido el escenario de constante toma de decisiones autónomas en relación con el tratamiento de los datos de carácter personal. No era mucho mejor el escenario que ofrecían ámbitos del sector público como el de la salud. Ciertamente, el rígido marco definido por la legislación sobre investigación biomédica, y la regulación de los ensayos clínicos y de los estudios observacionales, vino a traer algo de disciplina en este ámbito. Sin embargo, más allá de lo dispuesto en el Código Tipo de Farmaindustria, lo más usual era que la investigación en salud con datos se resolviera favorablemente por los comités de ética hospitalaria de la investigación con el expediente de la mera presentación de un modelo de consentimiento informado.

Esta cultura aún permanece en nuestros días. En muchas de nuestras instituciones, lejos de haber diseñado sistemas robustos, ordenados a garantizar el cumplimiento normativo, siguen existiendo una suerte de organización en taifas en las que tanto la gestión como la investigación no sólo se despliegan en los sistemas de información gestionados por los tradicionalmente conocidos como servicios de informática, sino por una galaxia de servidores comprados con fondos de investigación o mediante contratos menores. Ello cuando no son directamente programados sin supervisión jurídica, y sin ajuste a las políticas de seguridad corporativas, por el propio investigador o por terceras empresas o becarios. Cuando el desconocimiento o las capacidades quedan lejanas de aquello que exige el Esquema Nacional de Seguridad y la normativa se está invirtiendo en una catástrofe segura.

Únase a ello que, con muy pocas excepciones, y al igual que sucede con el conjunto de la Administración Pública española, la formación en protección de datos es puramente voluntaria, no obedece a un esfuerzo sistemático, estratégico y ordenado y, por tanto, no alcanza al conjunto de la organización. Así, el conocimiento de una normativa estratégica para el futuro de nuestra transformación digital se distribuye de forma asimétrica, sin los niveles de profundidad y especialidad adecuados, y sin que en ninguna circunstancia el personal se sienta preocupado y comprometido con los valores de su organización. Ello sin perjuicio de un desconocimiento en ocasiones palmario de las obligaciones de seguridad.

Basta con proponer una apuesta al lector para entender de modo muy directo a qué nos referimos. ¿Recuerda haberse registrado en un evento universitario en un formulario de un proveedor privado no autorizado por la institución universitaria y bajo la única autoridad de un correo electrónico que ni siquiera pertenecía al dominio de la universidad? Tampoco se sorprenda de que algún delegado de protección de datos considere razonable que el personal utilice cuentas privadas provistas por el operador dominante sin la existencia de ningún contrato que garantice las obligaciones de un encargado del tratamiento, sencillamente porque resultan prácticas usuales en el contexto de su universidad. Nos reconviene al lector, ya que hemos descendido a lo más anecdótico, dejando de lado riesgos más severos como la pérdida de datos, la utilización de datos procedentes de otros sistemas de información, como el sanitario por parte del profesorado con plaza vinculada, o al desarrollo de software y aplicaciones móviles sin la menor aplicación de los criterios de protección de datos desde el diseño y por defecto. Cuestión esta última, no solo presente en el caso de decisiones autónomas por parte del personal, sino también en muchas ocasiones en la contratación de servicios por parte de la universidad de un extraordinario riesgo y pese a la existencia de informes claros, precisos y muy contundentes por parte del regulador respecto de los riesgos en los que se podía incurrir.

No es esta la Universidad de la que nos gusta hablar ni tampoco un dibujo fiel del momento en el que nos encontramos. Un momento, sin duda, en fase de transición, en el que después de lustros de ímprobo esfuerzo comienzan a verse los primeros resultados derivados del esfuerzo colectivo en las Universidades Españolas y Crue de los profesionales de las tecnologías de la información, de la privacidad y la seguridad y de un trabajo individual en el que nunca han cejado ni renunciado en su lucha por alcanzar la excelencia. Sin embargo, es deber del investigador y del experto no esconder aquello que se conoce y que a su juicio constituye un gravísimo riesgo para la transformación digital de la institución universitaria.

La transformación a la que nos venimos refiriendo, la generación de repositorios de datos de investigación, y el acceso como usuarios a recursos de terceros está siendo abordado por un conjunto de regulaciones a las cuales nos vamos a referir, aunque sin entrar en el detalle. En primer lugar, cabe referirse a la llamada Data Governance Act. Esta norma busca que tanto en el sector público como en el privado surjan servicios de intermediación en el mercado de los datos, con capacidad de ofrecer servicios de procesamiento que puedan superar los riesgos derivados del tratamiento de datos de carácter personal y la garantía de derechos de las entidades titulares de los conjuntos de datos, como los de propiedad intelectual y secreto industrial y comercial. Ello implica un diseño de ecosistemas del tratamiento altamente dependiente de un trabajo bien hecho en el desarrollo de las correspondientes evaluaciones de impacto en la protección de datos y de la adopción de un conjunto de decisiones que deben buscar la excelencia. Para ello, se acompaña en los modelos más maduros un conjunto de decisiones relacionadas con las condiciones de diseño del software, del establecimiento de modelos de gobernanza jurídica tecnológica y ética y de la aplicación de los criterios más avanzados en materia de análisis de riesgos en el uso de la analítica de datos o la inteligencia artificial.

Otro tanto sucede con la Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre el Espacio Europeo de Datos Sanitarios. Si nos molestamos en consultar su estado actual de tramitación, se aprecia que en ámbitos como los requisitos que deberán cumplirse en su momento para solicitar el acceso a la información a través de un organismo de acceso a datos, o las mínimas medidas de seguridad que se definen para los sistemas de información, obligan a reconsiderar de modo muy riguroso la estructura actual y el modelo de gestión de la información para la investigación universitaria. Por otra parte, algunos aspectos de esta regulación, como los requerimientos para el desarrollo de software de historia clínica o de aplicaciones de bienestar, la programación de las condiciones de control remoto por parte de las de los pacientes respecto de sus datos, o la apelación a esquemas de certificación, constituyen un claro toque de atención: las cosas deben cambiar.

Bajo los mismos principios se desarrolla la Propuesta de Reglamento de inteligencia artificial. Una norma que se basa en tres pilares esenciales: un enfoque centrado en el riesgo de los sistemas de información, un diseño capaz de garantizar el pleno respeto a la dignidad humana y a los derechos fundamentales y, finalmente, un enfoque centrado en el ser humano, tanto desde la perspectiva de la garantía de su dignidad como desde el punto de vista de la libre decisión en cuanto al uso de estos sistemas y el control final sobre los mismos.

Asistimos, por tanto, a una revolución para la cual el cumplimiento normativo y el compromiso con el mismo por parte de la organización universitaria es crucial. En primer lugar, porque ni los profesionales en la institución ni los futuros egresados pueden permitirse el lujo de operar en el mercado sin un conocimiento profundo de la materia y sin un empoderamiento que les haga comprometerse plenamente con su garantía y respeto. En segundo lugar, porque a poco que nuestras infraestructuras sean sometidas al estrés regulador seguramente estallarán por todas sus costuras. Y

una Universidad en la que no sea viable el cumplimiento de todas y cada una de las normas a las que nos acabamos de referir, será una paria en la Unión Europea de los espacios de datos. Cuando solicite el acceso a información a través de un organismo de acceso a datos se arriesgará a un rechazo cuando no cumpla con los requisitos mínimos exigibles, o sencillamente a ser sancionada cuando acabe cometiendo una infracción. Asimismo, aquella universidad que no sea capaz de exhibir músculo, de ofrecer confianza al mercado, no podrá competir en la captación de fondos privados.

Por otra parte, y se habla desde la experiencia, en el momento en el que, en proyectos de investigación financiados por la Unión Europea, los auditores y revisores encuentren el menor error, corremos el riesgo de una pérdida significativa de los recursos disponibles. Es obvio que un proyecto de investigación con datos, financiado por la Unión que no reúna los requisitos éticos y jurídicos que conforman paquetes de trabajo impuesto de manera obligatoria como requisito imprescindible para la investigación con datos, podrá ser excluido de la financiación.

Sin embargo, existe un riesgo más relevante. Una organización que no sea capaz de cumplir con el bloque normativo que ordena el uso de los datos personales y no personales, no podrá generar conjuntos de datos de alta calidad. Porque cuando existan dudas en cuanto a la legalidad de un tratamiento, cuando se proyectan modelos de tratamiento anárquicos que no permiten asegurar la actualización y fiabilidad de los datos, cuando constantemente se transita el delgado filo de la navaja en materia de seguridad, uno no puede asegurar que las decisiones basadas en sus propios datos puedan ser mínimamente confiables.

Desgraciadamente, en este artículo hemos decidido emplear un tono severo y admonitorio. Hubiera sido más sencillo regalar los oídos y la mente de nuestros lectores con el escenario idílico del futuro prometedor que nos ofrece la transformación digital. Y no hubiéramos mentido. Sin embargo, les hubiera estado proponiendo que, al igual que Moisés, nos quedemos a las puertas de la tierra prometida de Canaán sin poder entrar en ella. Cada minuto que transcurre podríamos estar perdiendo la valiosa oportunidad de tomar decisiones estratégicas para nuestra universidad. De afrontar una estrategia de reingeniería de procesos que mejore nuestra eficiencia en la gestión de la información, asegure una mejor administración y libere a nuestro personal de esfuerzos innecesarios.

Debemos afrontar, también, decisiones cruciales en torno a nuestras infraestructuras de procesamiento de datos en un mundo que está migrando rápidamente a entornos de cloud o a una federalización de sus esfuerzos mediante espacios federados de datos. Lo contrario, el reducido mundo del proyecto de investigación, del departamento o de la facultad, está sin duda condenado al fracaso.

Estamos todavía a tiempo de reforzar nuestras plantillas con recursos humanos un alto conocimiento en materia de derecho digital, de protección de datos, de seguridad de la información, de anonimización y de ética de la inteligencia artificial. Si nos empeñamos en una concepción del personal como administración de carácter general, contratada mediante largos procesos burocráticos y en condiciones que no pueden competir con el mercado, nos veremos privados de la creación de un cuerpo de analistas de datos y expertos en computación, que resulta vital para nuestro futuro. La tarea es ingente, pero alcanzable. Es la hora de que aquellos que asumen el liderazgo de nuestras organizaciones afronten un proceso de transformación digital que debería empezar por la propia transformación de nuestras mentes, y de nuestro modo de ver el mundo, la investigación y la gestión.

4. CERTIFICACIONES EN EL ÁMBITO DE LOS DATOS

Ismael Caballero

Universidad de Castilla-La Mancha



Cada vez son más las organizaciones que son conscientes de la necesidad de invertir una serie de esfuerzos en asegurar la calidad de sus activos de datos. Si se consigue este aseguramiento, será posible transmitir confianza en que, del uso adecuado de estos activos de datos se podrá extraer el máximo valor de ellos. La necesidad de estos esfuerzos se ha acentuado incluso más con la adopción de tecnologías de determinadas tecnologías disruptivas tales como Big Data, Analytics (incluyendo Deep Learning o Machine Learning) o Internet of Things (IoT) por citar algunas de ellas.

Es posible definir tres vías para realizar estos esfuerzos:

- A través de la gestión de la calidad de los repositorios de datos
- A través de la implementación e institucionalización de buenas prácticas relacionadas con la gestión de la calidad de los datos; y como estas prácticas están íntimamente relacionadas con las buenas prácticas de gestión de datos (desde el punto de vista puramente tecnológico) y desde el punto de vista del gobierno de datos (desde el punto de vista puramente organizativo, resulta imprescindible combinar la implementación de todos estos tipos de buenas prácticas
- A través de la formación y capacitación de las personas que desempeñan las responsabilidades relacionadas con las buenas prácticas anteriores

En numerosas ocasiones, y aunque las organizaciones trabajen en todas o algunas de estas vías, puede resultar interesante refrendar los resultados obtenidos mediante un reconocimiento explícito basado en la inspección de evidencias objetivas capturadas de forma rigurosa y libre de intereses por parte de un tercero independiente, que usará modelos de referencia desarrollados y respetados internacionalmente. Este reconocimiento se obtiene en forma de certificaciones. En este documento se van a explicar algunos de los marcos de certificación más relevantes existentes hoy en día en el dominio de los datos con respecto a:

- Certificación del nivel de calidad de repositorios de datos
- Certificación del nivel de madurez de las buenas prácticas de gestión de datos, gestión de calidad de datos y gobierno de datos
- Certificación del nivel de conocimiento y habilidades de los profesionales que desempeñan las responsabilidades relacionadas con los datos

Es comúnmente aceptado que un conjunto de datos tiene calidad cuando estos sirven para el propósito para el que fueron recogidos (fitness for use) (Strong et al., 1997). Para evaluar la calidad de los datos es necesario identificar un conjunto de criterios que permitan representar los requisitos de calidad de datos de los usuarios. A estos criterios se les llama dimensiones o características de calidad de datos. Al conjunto de dimensiones o características de calidad de datos se les llama modelo de calidad de datos (Wang, 1998). Es necesario poder medir estas características mediante la definición de unos procedimientos de medición, típicamente basados en las reglas de negocio que regulan la adecuación al uso de los datos en un contexto determinado (Loshin, 2002). Una vez que se tienen estas mediciones, es necesario comparar los resultados de dichas mediciones con los valores mínimos deseables especificados por los usuarios de los datos a través de los requisitos de calidad de datos, y que dependen fuertemente del apetito de riesgo de la organización a la hora de utilizar estos datos (Redman, 2016). Diferentes autores han proporcionado diferentes mecanismos para medir la calidad de los datos de acuerdo con las dimensiones de calidad seleccionadas y se han propuesto numerosas metodologías de evaluación (Batini & Scannapieco, 2016). El principal problema de estas contribuciones es que típicamente se han realizado ad hoc y no permitían ni generalizar los resultados obtenidos ni compararlos con los obtenidos por otras organizaciones (Loshin, 2011). A fin de paliar estos problemas, algunas organizaciones internacionales han realizado contribuciones intentando estandarizar los conocimientos y principios básicos comunes para medir y evaluar la calidad de los datos. Algunos de estos esfuerzos se han materializado en estándares internacionales como ISO/IEC 25012 (ISO/IEC, 2008a) o como ISO 8000-8 (ISO, 2015), que habilitan la posibilidad de obtener una certificación de cumplimiento.

Para poder utilizarlos, es necesario desarrollar determinados aspectos de estos estándares, lo que requiere una particularización de los marcos propuestos; no obstante, estas particularizaciones se pueden hacer sobre una base común en la que han participado numerosos expertos de numerosos países provenientes de diferentes industrias, lo que garantiza unas líneas comunes de trabajo.

En este capítulo se presenta un esquema de certificación del nivel de calidad del repositorio de datos desarrollado conjuntamente por el Área de Certificaciones TIC de AENOR Internacional y AQCLab, una spin-off de la Universidad de Castilla-La Mancha. Es importante mencionar que AQC Lab es la única organización a nivel internacional que ha sido acreditada por ENAC para la elaboración de informes de evaluación del nivel de calidad de datos de repositorios de datos.

El esquema de certificación consta de un modelo de calidad de datos y una metodología de evaluación.



Modelo de Calidad de Datos

El **modelo de calidad de datos** está basado en ISO/IEC 25012. Este estándar introduce y define quince características de calidad de datos (véase Tabla 1) cuya medición y evaluación se puede realizar desde dos puntos de vistas:

- Calidad de datos inherente se refiere a la medida en que las características de calidad de los datos tienen el potencial intrínseco de satisfacer las necesidades declaradas e implícitas cuando los datos se utilizan en condiciones específicas. Por ejemplo: valores del dominio de datos y posibles restricciones (reglas de negocio que rigen la calidad requerida para la característica en una aplicación determinada, relaciones entre los valores de datos o con los metadatos...
- Calidad de datos dependiente del sistema se refiere a la medida en que se alcanza y preserva la calidad de los datos dentro de un sistema informático cuando los datos se utilizan en condiciones específicas. Desde este punto de vista, la calidad de los datos depende del dominio tecnológico en el que se utilizan los datos; se logra mediante las capacidades de los componentes de los sistemas informáticos, tales como: dispositivos de hardware (por ejemplo, para hacer que los datos estén disponibles o para obtener la Exactitud requerida), software de sistemas informáticos (por ejemplo, software de copia de seguridad para lograr recuperabilidad) y otro software (por ejemplo, herramientas de migración para lograr la portabilidad).

Cabe mencionar que algunas de las características pueden medirse/evaluarse desde los dos puntos de vistas. En cualquier caso, las características que son incluidas en el modelo de calidad certificable son las consideradas como puramente inherentes: exactitud, completitud, consistencia, credibilidad y actualidad.

Tabla 1. Características de Calidad de Datos definidas en ISO/IEC 25012 (ISO/IEC, 2008b).

Características	Definición	Inherente	Dependiente del sistema
Exactitud	Grado en que los datos tienen atributos que representan correctamente el valor de un atributo	●	
Complejidad	Grado en el que existen suficientes valores para todos los atributos necesarios para la representación de una entidad	●	
Consistencia	Grado en que los datos están libres de contradicción y son coherentes con el resto de los datos de su contexto de uso	●	
Credibilidad	Grado en que los datos se consideran ciertos y creíbles por los usuarios	●	
Actualidad	Grado en que los datos tienen atributos con las fechas y tiempos correctos	●	●
Accesibilidad	Grado en que los datos pueden ser accedidos por cualquier usuario	●	●
Cumplimiento	Grado en el que los datos están contruidos conforme estándares, convenciones o regulaciones	●	●
Confidencialidad	Grado en el que los datos tienen atributos específicos que solo pueden ser accedidos por usuarios autorizados	●	●
Eficiencia	Grado en el que los datos tienen atributos que pueden ser procesados y provistos dentro de los niveles de rendimiento esperados	●	●
Exactitud	Grado en el que los datos tienen atributos que son precisos	●	●

Trazabilidad	Grado en el que los datos tienen atributos que proveen información detallada sobre los cambios realizados en los datos	●	●
Comprensibilidad	Grado en el que los datos se expresan de manera que los usuarios puedan leerlos e interpretarlos correctamente	●	●
Disponibilidad	Grado en el que los datos están disponibles para ser accedidos por usuarios y/o aplicaciones autorizadas		●
Portabilidad	Grado en el que los datos pueden ser alojados, reemplazados o movidos desde un sistema a otro		●
Recuperabilidad	Grado en el que los datos disponen de formas de mantener un nivel especificado de operabilidad incluso cuando se producen fallos		●

Para medir cada una de estas características, se toma de ISO/IEC 25024 (ISO/IEC, 2015) la idea de propiedad de calidad de datos, que representa un aspecto medible de cada característica (véase Figura 5). Esto permite realizar estudios detallados de las mediciones de estas propiedades para profundizar en el análisis de las causas raíz de las razones de los niveles inadecuados de calidad de datos. Estos estudios habilitarán poder definir de forma más precisa y eficiente las posibles acciones posteriores de limpieza de datos – en el caso de que los resultados de la evaluación no hayan sido los deseados-. Por ejemplo, la evaluación de la exactitud implicaría la medición de las tres siguientes propiedades: exactitud semántica, exactitud semántica y rango de exactitud.

FIGURA 5. Propiedades medibles de las características de calidad de datos extraídas de ISO/IEC 25024

Exactitud	Compleitud	Consistencia	Credibilidad	Actualidad
Exactitud Sintáctica	Compleitud de Registro	Integridad Referencial	Credibilidad de la fuente	Frecuencia de Actualización
Exactitud Semántica	Compleitud de Fichero	Consistencia de Formato	Credibilidad de los valores de datos	Conveniencia de Actualización
Rango de Exactitud	Compleitud de Valores de Datos	Riesgo de Inconsistencia		
	Falsa Compleitud de Fichero	Consistencia Semántica		

Las medidas (función de medición) correspondientes a las propiedades son también extraídas de ISO/IEC 25024 (véase Figura 6). Estas funciones de medición –típicamente en formato densidad- requieren identificar, listar y validar con los principales implicados en la evaluación de la calidad de los datos una serie de reglas de negocio que rigen la adecuación de los datos al uso que se pretende dar.

FIGURA 6. Ejemplos de definición de la propiedad «Exactitud Sintáctica de datos» de la característica «Exactitud» según ISO/IEC 25024

ID	Nombre	Descripción	Función de medición	Partes de ciclo de vida. Entidades objetivos. Propiedades
Acc-I-1	Exactitud sintáctica de datos	Relación de cercanía de los valores de los datos a un conjunto de valores definidos en un dominio	$X=A/B$ A=números de elementos de datos que tienen valores relacionados sintácticamente exactos B=números de elementos para los que se requiere una exactitud sintáctica	Todas excepto diseño de datos Fichero de datos ítem de datos y valor de datos

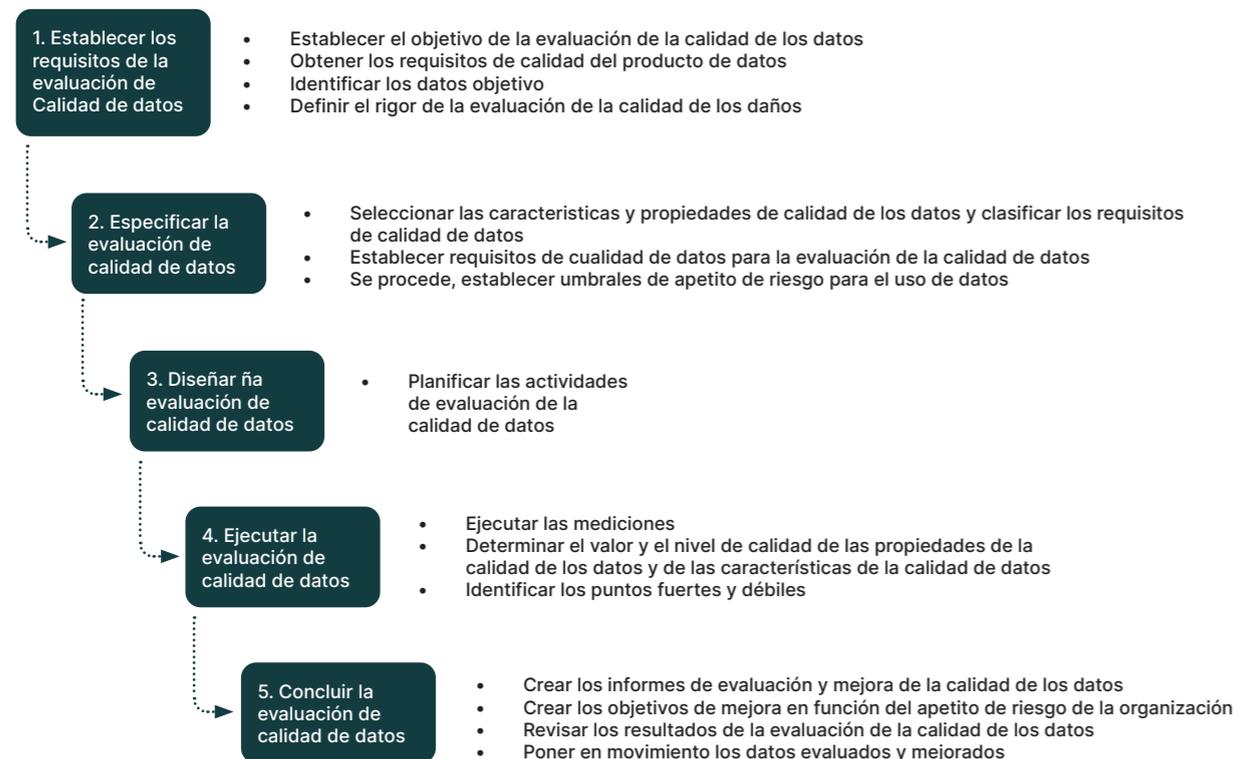
Una vez consensado el listado de reglas de negocio, y como no todas son susceptibles de ser usadas para la evaluación de todas las características de calidad de datos, estas reglas de negocio se deben clasificar o agrupar primero para cada característica seleccionada y, luego, para cada una de las propiedades. La tarea de identificación de reglas de negocio implica la inspección de la documentación de los sistemas de información existentes en la organización (incluyendo software y bases de datos) y entrevistas con los principales implicados, y un fuerte proceso de refinado y limpieza de los datos. Es importante acortar y negociar con los principales implicados el resultado final del listado y clasificación de las reglas de negocio porque de ello depende fuertemente los resultados de la evaluación: un solo cambio en el listado de reglas de negocio implica obtener diferentes resultados en la evaluación. Como parte del modelo de evaluación definido conjuntamente por AENOR y AQCLab, se han desarrollado funciones de agregación (llamadas funciones por perfiles) que permiten combinar los resultados de la medición de las propiedades para generar un valor del nivel de calidad de datos para cada una de las características. Estos valores de calidad están reglados entre 1 y 5, siendo 1 el mínimo valor de calidad y 5 el máximo. AENOR otorga la certificación si todas las características seleccionadas alcanzan al menos un nivel de 3.



Metodología de evaluación de calidad de datos

Una metodología de evaluación de calidad de datos de los repositorios de datos basada en ISO/IEC 25040. Esta metodología permite abordar de forma rigurosa y sistemática el proceso de evaluación y medición de calidad de datos en los repositorios seleccionados. Esta metodología (véase Figura 7) consta de cinco tareas principales con una serie de tareas bien definidas y acotadas.

FIGURA 7. Metodología de evaluación basada en ISO/IEC 25040



En el proceso de certificación participan los siguientes actores:

- Entidades privadas o instituciones de la Administración Pública (como universidades) que quieren certificar el nivel de calidad de datos de sus repositorios de datos con respecto a una o varias características de calidad inherentes del modelo presentado
- AQC Lab, que elabora informes de evaluación para las características de calidad seleccionadas teniendo en cuenta las reglas de negocio para la medición de las propiedades medibles de las características de calidad de datos seleccionadas. Dichos informes son las entradas que AENOR toma como entrada para decidir si otorga una certificación a un repositorio de datos
- AENOR Internacional, que es la organización que decide y gestiona, en base a los resultados de los informes acreditados, si concede el certificado correspondiente (véase Figura 8 para un ejemplo de certificación de un repositorio de datos concedido a una empresa española de viajes)

FIGURA 8. Ejemplo de Certificado de calidad de datos de un repositorio de datos



En el artículo (Gualo et al., 2021) se explica el modelo con todos los detalles posibles y se presentan experiencias de proyectos industriales de certificación.

Certificación del nivel de madurez organizacional con respecto al gobierno del dato, la gestión del dato y la gestión de la calidad del dato

Una certificación del nivel de calidad de un repositorio de datos solo acredita el estado del repositorio de datos en un momento dado. Para garantizar que los niveles de calidad alcanzados -que incluso han podido ser certificados- no se degradan en el tiempo tan pronto como el repositorio de datos entre en producción, la organización debería adoptar e implantar una serie de buenas prácticas de gestión de calidad de datos.

La implantación de estas buenas prácticas -relacionadas con la medición, evaluación y limpieza de los datos- hace que se pueda trabajar de forma sistemática, rigurosa, predecible y sostenible. Pero estas buenas prácticas no deberían considerarse de forma aislada, sino que dependen fuertemente de otras prácticas de gobierno de datos (que deben transmitir una dirección sobre cómo implementar la estrategia de datos de la organización) y con otras prácticas de gestión de datos (entendida desde el punto de vista de gestión de la infraestructura tecnológica para dar soporte al uso de los datos por parte de los procesos de negocio; esta infraestructura tecnológica también debe dar soporte a las prácticas de gestión de calidad de datos).

En función de cómo la organización haya implementado, institucionalizado y siga estas buenas prácticas es posible hablar de madurez de la organización con respecto a estas buenas prácticas (o procesos) de datos.

Se puede describir la madurez en términos de capacidad (excelencia en la gestión que habilita un potencial para realizar ciertas tareas con garantías de éxito), en términos de realización de determinadas tareas (se pueden organizar en orden de dificultad creciente ciertas tareas de modo que la organización puede estar más o menos preparada), o incluso en términos de determinados valores umbrales para determinados indicadores (calculados o medido a partir de ciertos aspectos observables de la organización o de las buenas prácticas durante la ejecución de la tarea). Diferentes organizaciones han presentado diferentes modelos de madurez organizacional (algunos de ellos admiten incluso una posible certificación) tal como el Data Maturity Model (DMM) (DMM, 2013) -actualmente en propiedad de ISACA-, Data Capability Model (DCAM) (EDM Council, 2020) o el Modelo Alarcos de Mejora de Datos (MAMD) (DQTeam, 2020) -desarrollo de la UCLM en colaboración con dos de sus spin-off, AQCLab y DQTeam-, todas con una aproximación a proceso.

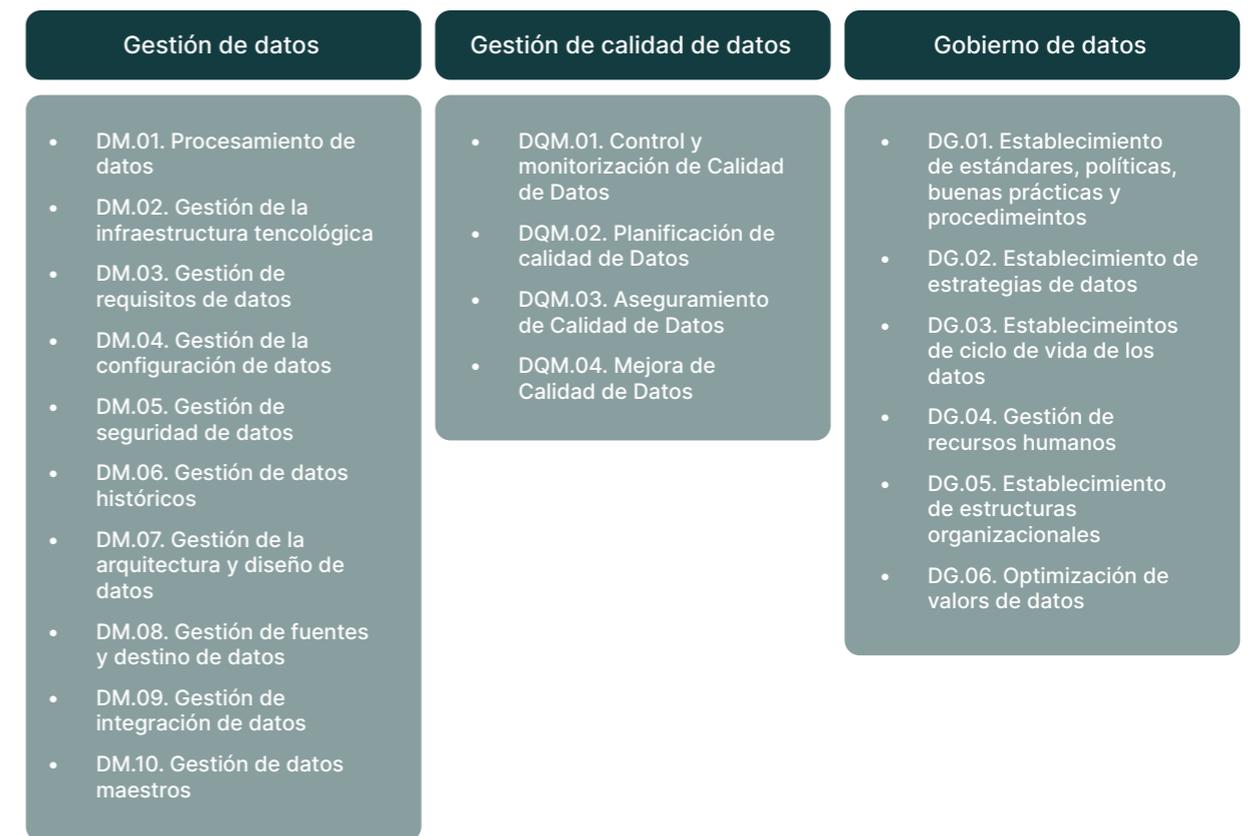
Estos modelos tienen sus ventajas (p.ej. planteamiento de diferentes filosofías para abordar la forma de implantar las buenas prácticas usando sus correspondientes modelos de referencias de procesos) y sus desventajas (p.ej. importantes costes de licenciamiento de algunos de sus modelos). De entre todos los modelos, haremos especial hincapié en MAMDv3.0 por varias razones: (1) está basado en diferentes estándares internacionales abiertos; (2) es certificable por AENOR Internacional en el ámbito internacional con costes razonables; (3) es fácilmente aplicable en organizaciones de diferentes tamaños, y (4) no tiene costes de licenciamiento ya que puede ser conseguida sin costes de la página web de DQTeam (<https://mamd.dqteam.es>).

Modelo Alarcos de Mejora de Datos MAMDv3.0

MAMD es un modelo para la evaluación y mejora de la madurez organizacional de los procesos de datos (gestión de datos, gestión de calidad de datos y gobierno de datos). MAMD consta de dos componentes principales:

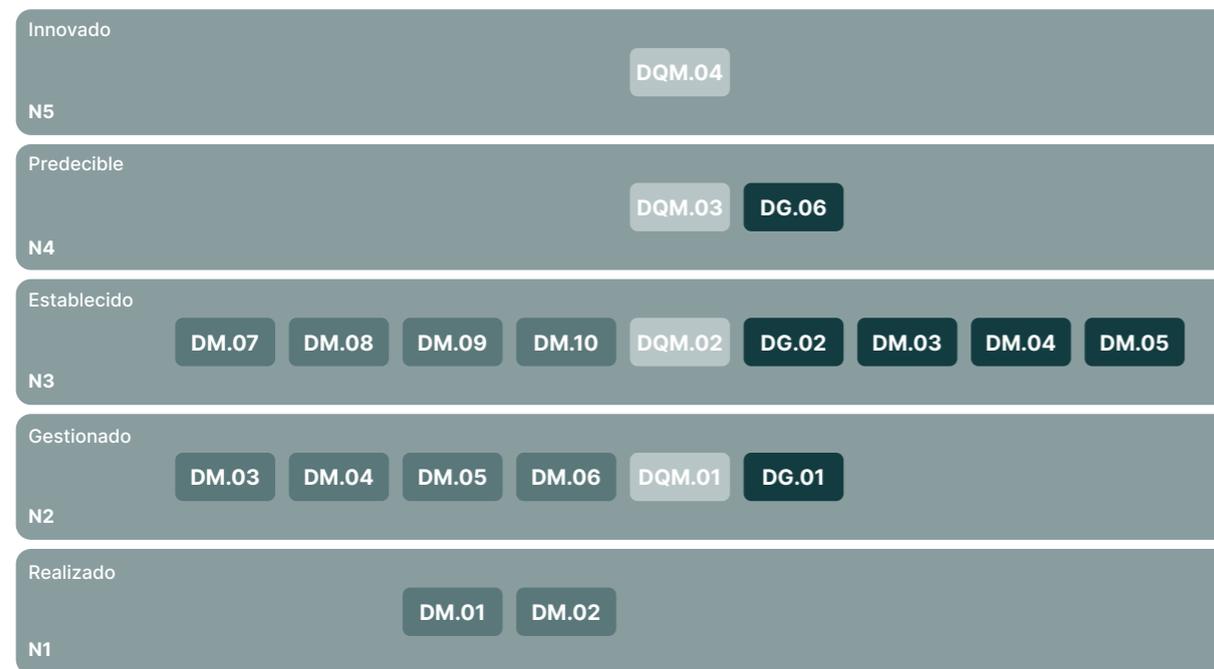
- Un **modelo de referencia de procesos (MRP)** basado en el modelo de referencia propuesto en ISO/IEC 8000-61 (ISO, 2016) (cuyo principal enfoque es la gestión de calidad de datos). Este modelo fue extendido con los principios de procesos de gobierno (DG) establecidos en ISO/IEC 38505-1 (ISO/IEC, 2017) e ISO/IEC 38505-2 (ISO, 2018b) y enriquecido con las buenas prácticas de gestión de datos y de gobierno de datos proveniente de otros marcos de referencias en el sector, tales como el mencionado DMM, u otros tales como COBIT2019 (ISACA, 2019) o el DMBok (DAMA, 2017) de DAMA. El resultado es un conjunto de tres grupos de procesos (véase Figura 9), que comprende 10 procesos de gestión de datos, 4 procesos de gestión de calidad de datos, y 6 procesos de gobierno de datos. Cada uno de estos procesos está descrito por medio de cuatro aspectos: propósito, resultados de procesos, tareas y productos de trabajo

FIGURA 9. Modelo de referencia de procesos de MAMDv3.0



- Un **modelo de evaluación de procesos (MEP)** basado en distintas partes de ISO 33000, del que se toman prestados algunos elementos tales como una metodología de evaluación, o los conceptos de capacidad y modelo de madurez entre otros; el modelo de madurez tiene dos dimensiones: dimensión de capacidad de proceso -que especifica diferentes aspectos de excelencia en la gestión establecidos mediante cinco niveles de capacidad descritos mediante una serie de requisitos para cada uno de los niveles- y una dimensión de procesos -que organiza los procesos del modelo de referencia de procesos ordenándolos por niveles creciente de dificultad-. Este modelo de madurez está basado en el modelo de madurez propuesto en ISO 8000-62 (ISO, 2018a) y reorganiza los procesos nuevos añadidos en el modelo de referencia de procesos. El modelo de madurez se representa en la Figura 10

FIGURA 10. Modelo de madurez incluido en el modelo de evaluación de procesos de MAMDv3.0



A la hora de evaluar el nivel de madurez organizacional (evaluación basada en la fortaleza de las evidencias recogidas a través de un proceso de auditoría que inspecciona el grado de implantación de las buenas prácticas a través de la inspección de un conjunto de procesos de negocio), las organizaciones, en función de sus necesidades y prioridades, fijarán el alcance de la evaluación, deseablemente con respecto a un nivel de madurez específico. Si no han alcanzado el nivel objetivo, se les proponen acciones de mejora para cada uno de los procesos evaluados. Si lo alcanzan, y están interesadas en certificarse, entonces, AENOR INTERNACIONAL puede gestionar el proceso de certificación. De entre las organizaciones que han adoptado este modelo, la Biblioteca General de la UCLM, ha sido la primera en obtener la certificación (<https://revista.aenor.com/340/nueva-certificacion-para-el-gobierno-gestion-y-calidad-de-lo.html>) (véase Figura 11).

FIGURA 11. Ejemplo de Certificado de Metodología de evaluación basada en ISO/IEC 25040



42 <https://cdmp.info/>

Las certificaciones de competencias personales tienen como objetivo acreditar que un profesional ha adquirido una serie de conocimientos, habilidades o destrezas con los datos para desarrollar distintos tipos de trabajo, tanto en el ámbito de la gestión de los datos como en el del gobierno de los datos. Tener a profesionales certificados proporciona una certidumbre a la organización de que los activos de datos están siendo gestionados y gobernados adecuadamente. Se pueden distinguir varios tipos de certificaciones profesionales:

- Certificaciones que reconocen que un profesional ha alcanzado un determinado grado de madurez en el uso de ciertas tecnologías. Estas certificaciones son gestionadas y expedidas normalmente por los fabricantes y vendedores de dichas tecnologías. Por ejemplo, Oracle, Microsoft o Anjana, por citar algunos ejemplos, tienen sus propias certificaciones que acreditan que los profesionales pueden tener un buen desempeño usando las tecnologías de gestión y gobierno de datos
- Certificaciones que reconocen que un profesional ha alcanzado un determinado grado de madurez en los conocimientos y habilidades de gestión de datos, gestión de calidad de datos y gobierno de datos. Es decir, están más orientadas a que los profesionales puedan demostrar que han adquirido estas competencias. A nivel nacional, la oferta/demanda de estas certificaciones está cubierta por las dos siguientes:

43 <https://clubcdo.com/certged/>

<https://www.ie.edu/es/lifelong-learning/programas/programa-certged-certificacion-en-gobierno-y-estrategia-del-dato/>

44 <https://cdmp.info/about/>

- La «**certificación en Gobierno y Estrategia del Datos**» (CertGed) gestionada en colaboración con Club CDO Spain e IE Business School⁴³. En esta certificación, los candidatos adquieren conocimientos sobre gobierno de información, cadena de valor del dato, seguridad y privacidad, y estrategia del dato.

Las organizaciones certificadoras suelen establecer varios niveles de certificación de competencias profesionales en función de la profundidad de conocimientos requeridas y/o de los requisitos que deben satisfacer los profesionales para obtener dichas certificaciones. Esto les habilita para el desempeño de diferentes funciones de distinta complejidad. Así, por ejemplo, DAMA establece los mostrados en la Figura 13.

En el caso de CDMP, es posible además realizar diferentes certificaciones para demostrar conocimientos en las siguientes áreas específicas del DMBOK-2, como pueden ser:

- Modelado y Diseño de Datos
- Metadatos
- Calidad de datos
- Gobierno de Datos
- Data Warehousing y Business Intelligence
- Gestión de Datos Maestros y Datos de Referencias
- Integración e interoperabilidad de datos

Figura 12. Marco de trabajo DMBOK-2 de DAMA Internacional (DAMA, 2017)

La certificación Certified Data Management Professional (CDMP) gestionada por DAMA International. Con esta certificación, el candidato demostrará el conocimiento y capacidad de aplicar el marco DMBOK de DAMA (véase Figura 12).

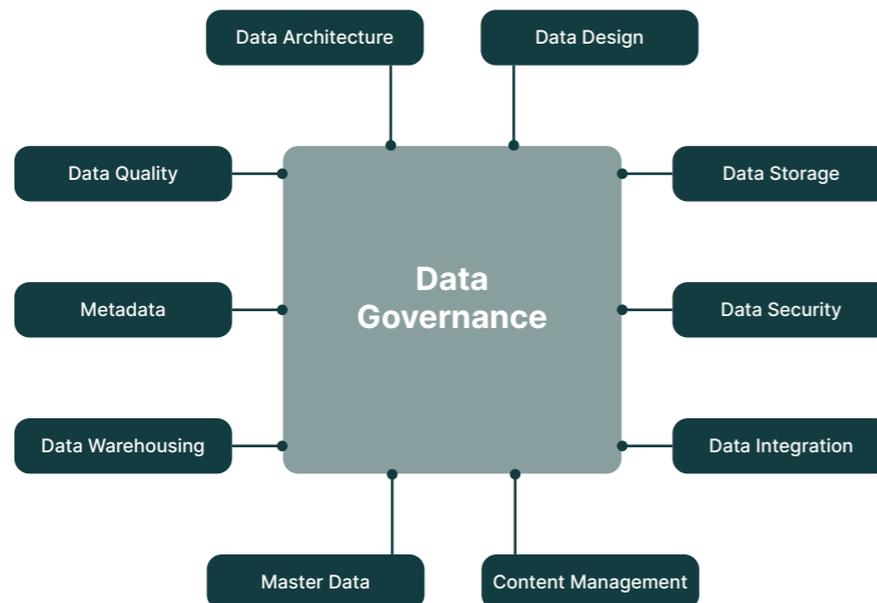


Figura 13. Niveles de certificación de CDMP y requisitos exigibles para obtener las certificaciones⁴⁴

CDMP Associate	CDMP Practitioner	CDMP Master	CDMP Fellow
Membership Central	Membership Central	Membership Central	Membership Central
Industry experience 6 month > 5 years	Industry experience 2 years - 10 years	Industry experience Minimum 10 years	Industry experience 25 years +
Requirements 60% pass Data Manag. Fundamentals exam	Requirements 70% pass Data Manag. Fundamentals exam 70% pass un 2 specialist exam	Requirements 80% pass Data Manag. Fundamentals exam 80% pass un 2 specialist exam	Requirements Globally recognised leadership Significant contribution to Data Manag. profession CDMP Master Contribution to CDMP & DMBDK By nomination

45 <https://www.uclm.es/estudios/proprios/curso-gobierno-gestion-datos>

En la actualidad, DAMA Spain colabora con distintas organizaciones en España para la impartición de cursos de estudio y preparación del examen de certificación de CDMP Fundamentals, como el título propio de la UCLM «Gestión y Gobierno de Datos»⁴⁵.

Este documento ha presentado un informe de las posibles certificaciones que permitirían transmitir confianza en que los datos que usa una organización son adecuados y tienen el valor suficiente para satisfacer los objetivos de las tareas para los que se están usando.

En cualquier caso, las organizaciones no se deberían plantear como objetivo el mero hecho de obtener dichas certificaciones, sino que deberían trabajar día a día para conseguir alcanzar los hitos marcados en las hojas de ruta, que, sin embargo, pueden ser demostrables mediante la certificación correspondiente.

Referencias

Batini, C., & Scannapieco, M. (2016). *Data and information quality*. Cham, Switzerland: Springer International Publishing.

DAMA. (2017). *DAMA-DMBOK: data management body of knowledge*. Technics Publications, LLC.

DMM. (2013). *CMMI Institute—Data Management Maturity (DMM)*. <https://stage.cmmi-institute.com/dmm>

DQTeam. (2020). *MAMD: Modelo Alarcos Mejora Datos v3,0*. <https://www.dqteam.es/index.php/mamd>

EDM Council. (2020). *The Data Capability Assessment Model (DCAM) Framework v2.2 Overview*. EDM Council. https://cdn.ymaws.com/edmcouncil.org/resource/collection/AC65DC50-5687-4942-9B53-3398C887A578/DCAM_Framework_v2_Overview_v2.2.1.pdf

Gualo, F., Rodriguez, M., Verdugo, J., Caballero, I., & Piattini, M. (2021). Data quality certification using ISO/IEC 25012: Industrial experiences. *Journal of Systems and Software*, 110938. <https://doi.org/10.1016/j.jss.2021.110938>

ISACA. (2019). *COBIT | Control Objectives for Information Technologies*. ISACA. <https://www.isaca.org/resources/cobit>

ISO. (2015). *ISO 8000-8:2015 Data quality – Part 8: Information and data quality: Concepts and measuring* [International Standard]. ISO.

ISO. (2016). *ISO 8000-61:2016*. ISO. <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/30/63086.html>

ISO. (2018a). *ISO 8000-62:2018: Information technology—Process assessment—Requirements for process reference, process assessment and maturity models*. ISO. <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/53/65340.html>

ISO. (2018b). *ISO/IEC TR 38505-2:2018 Information technology—Governance of IT—Governance of data—Part 2: Implications of ISO/IEC 38505-1 for data management*. *ISO/IEC TR 38505-2:2018 Information Technology — Governance of IT — Governance of Data — Part 2: Implications of ISO/IEC 38505-1 for Data Management*. <https://www.iso.org/standard/70911.html>

ISO/IEC. (2008a). *ISO/IEC 25012. Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model (Software Engineering — Software Product Quality Requirements and Evaluation (SQuaRE), p. 22* [International Standard]. ISO/IEC.

ISO/IEC. (2008b). *ISO/IEC 25012:2008 Software engineering—Software product Quality Requirements and Evaluation (SQuaRE)—Data quality model (Systems and Software Engineering -- Systems and Software Quality Requirements and Evaluation (SQuaRE))* [International Standard]. ISO/IEC.

ISO/IEC. (2015). *ISO/IEC 25024:2015 Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Measurement of data quality* [International Standard]. ISO/IEC.

ISO/IEC. (2017). *ISO/IEC 38505. Information technology—Governance of IT — Part 1: Application of ISO/IEC 38500 to the governance of data (Information technology — Governance of IT, p. 28)* [International Standard]. ISO/IEC.

Loshin, D. (2002). Rule-based data quality. 614-616. Scopus. <https://doi.org/10.1145/584792.584894>

Loshin, D. (Ed.). (2011). Preface. En *The Practitioner's Guide to Data Quality Improvement* (pp. xiii-xix). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-373717-5.00026-9>

Redman, T. C. (2016). *Getting in Front on Data: Who Does What*. Technics Publications.

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM*, 40(5), 103-110.

Wang, R. Y. (1998). A product perspective on total data quality management. *Communications of the ACM*, 41(2), 58-65.

5. LA UNIVERSIDAD EN LA ERA DEL DATO

Nuria Canal

Chief Data Officer Member



Desde mediados de los 90, estamos viviendo la llamada Cuarta Revolución Industrial (1995 –XXI) caracterizada por el imparable fenómeno de la digitalización de todos los sectores industriales, así como de la Administración pública, la educación y cualesquiera ámbitos sociales.

Esta revolución viene marcada por las últimas innovaciones en el ámbito de la robótica, la inteligencia artificial, IoT, Big Data, Blockchain, computación en la nube, 4G y 5G... toda una explosión de nuevas tecnologías con las que se pretende alcanzar la total independencia de la productividad de la mano de obra, así como la hiper-conectividad de los procesos productivos.

La tecnología y los sistemas, conectados y combinados entre sí, permiten poner en diálogo a las máquinas, dando lugar a una coordinación automatizada de los procesos y a la toma de decisiones descentralizada.

Algo que caracteriza y diferencia esta revolución de las anteriores, es la velocidad con la que tienen lugar los avances, que no tiene precedentes en la historia. Estamos acostumbrados a vivir en un estado continuo y permanente de cambio.

Se trata en definitiva de un cambio de paradigma, con grandes transformaciones en el ámbito de la conectividad, la movilidad, el uso de las energías, nuevos modelos de negocio, nuevos modelos de producción, nuevos modelos de educación, cambios culturales... que generan incertidumbre y debates sociales, éticos y políticos. La sociedad en su conjunto debe trabajar para que este cambio sea una oportunidad para todos.

El principal objetivo de una empresa u organización es captar clientes para ofrecerles sus productos y servicios, tratando, posteriormente, de fidelizarlos. Por ello, toda estrategia de transformación digital parte de la premisa de colocar al cliente en el centro de la organización, identificando sus necesidades y comportamientos, sus intereses, sus gustos, sus prioridades... para anticiparse ofreciéndole aquello que desean en su proceso de búsqueda y de compra.

Las tecnologías digitales convierten estos procesos en datos que, analizados, dan lugar a información y conocimiento empresarial con el que articular la toma de decisiones. Los datos nos hablan del negocio de una forma objetiva y dirigen la atención a lo más importante, nos indican pues el mejor camino a seguir para maximizar el negocio. Hablamos, por tanto, de organizaciones dirigidas por los datos, organizaciones data-driven.

Este mismo proceso, haciéndolo extensivo al conjunto del tejido empresarial y administrativo del país, da lugar a la denominada Economía del Dato, que considera el valor y riqueza generados por la explotación del dato en la economía en su conjunto.

La importancia de este reto es tal, que la UE ha creado los fondos Next Generation, los cuales reflejan claramente la voluntad por parte de la administración pública de apostar por los cambios que conlleva nuestra actual realidad digital.

La huella digital y la eclosión del Big Data

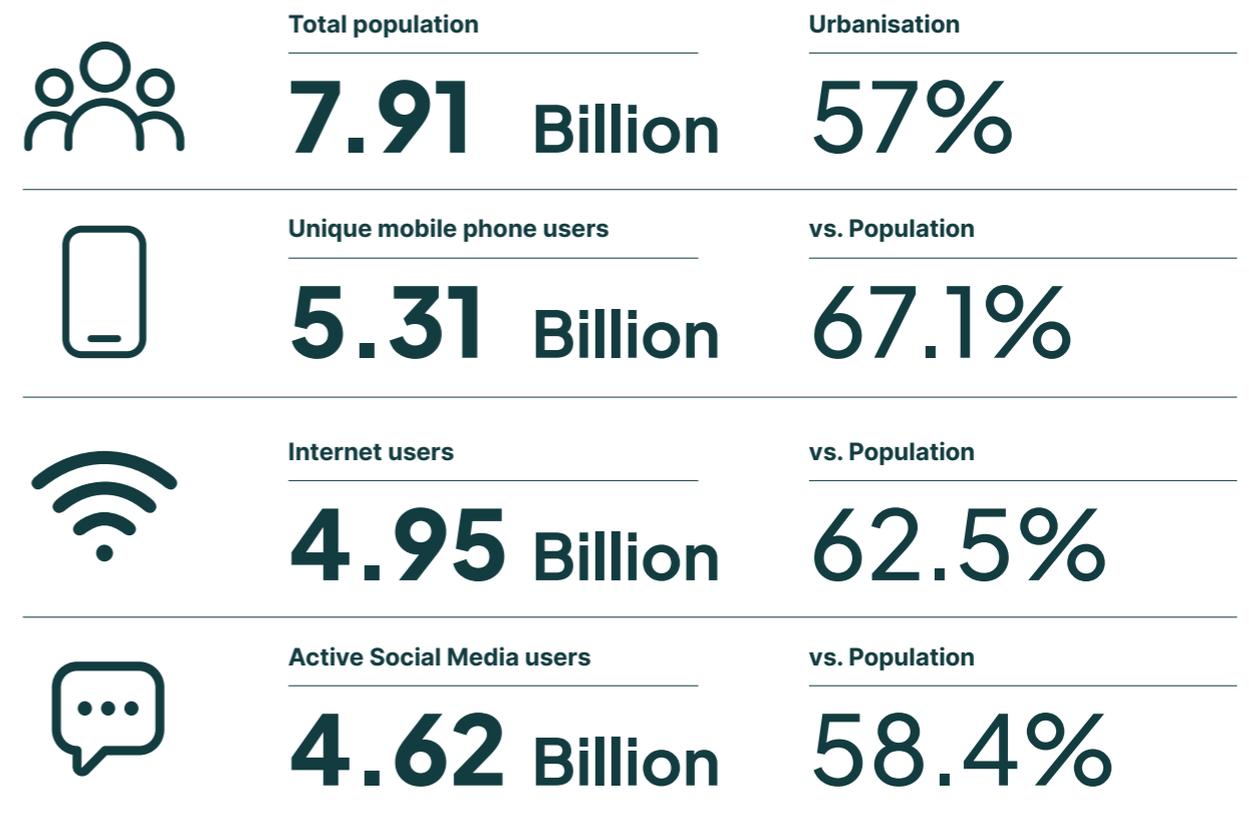
La alta penetración de Internet en los países –en los desarrollados supera el 90% y en el resto se alcanza el 50%- ha facilitado que la implantación, adopción y uso de las plataformas digitales y dispositivos inteligentes sea una realidad en nuestro día a día.

Según el estudio anual «Digital 2022. Global Digital Overview» de We Are Social, las dos regiones con mayor penetración de Internet en la actualidad son Europa oriental (94%) y el norte de Europa (98%), seguidas de cerca por Norteamérica (92%). Por otro lado, las regiones con menor penetración de Internet son África oriental (26%) y África occidental (42%).

Hay 4,62 billones de usuarios de social media, que supone un incremento del 10% vs 2021 y 5,31 billones de personas utilizando teléfonos móviles, que supone un incremento del 1,2% vs 2021.

Según el mismo estudio, en el año 2022, el número de internautas alcanza ya los 4,95 billones, lo que supone un +4% de incremento sobre enero del 2021, y que implica al 62,5% de la población mundial.

Figura 10. Essential digital headlines. Overview of the adoption and use of connected devices and services



En dicho estudio, se nos indica que actualmente una persona invierte en promedio 106 días al año conectado a la Red. En cada conexión se generan múltiples interacciones, cuyo registro genera la llamada huella digital; esto es, el conjunto de datos donde queda reflejado nuestro comportamiento, nuestros intereses y nuestras prioridades y que es, por tanto, de alto valor para todo tipo de empresa y organización.

Estos datos crecen de forma exponencial cada día, y es así como la capacidad de almacenar y procesar (Big Data) y analizar datos (Data Analytics), de forma masiva, se ha vuelto parte fundamental de las habilidades que han de desarrollar las organizaciones en su camino hacia la transformación digital.

El Big Data & Data Analytics aplicado a cualquier organización es, básicamente, el análisis masivo de sus datos para usarlos a la hora de tomar decisiones; por ello se han de adoptar sistemas capaces de analizar en cuestión de segundos, cantidades ingentes de datos que permitan responder a las demandas de la organización en tiempo real.

Los términos digitalización y transformación digital, se han vuelto habituales; aunque usados a menudo de forma indistinta, tienen en realidad marcadas diferencias.

La digitalización es el proceso de incorporar a la empresa, tecnologías digitales que sustituyan a las tradicionales para desempeñar la misma actividad, pero de modo más eficiente. Por ejemplo, la incorporación de un Customer Relationship Management, un Enterprise Resource Planning o un software para la gestión del personal interno (bajas, vacaciones, nóminas...), lo cual permita simplificar, mejorar y automatizar los procesos ya existentes. Dos de cada tres empresas españolas están, o ya han estado, en este proceso de renovación tecnológica, que les permite hacer sus procesos de forma mucho más rápida, eficiente y rentable.

Sin embargo, cuando hablamos de transformación digital, no solo pensamos en la introducción de nueva tecnología, sino también en la digitalización de los sistemas de trabajo de la mayoría de los departamentos de la empresa. Se trata, por tanto, de una transformación mucho más profunda, una reinención total de la empresa que requiere de la revisión y actualización del modelo de negocio, de sus productos y/o servicios, de su relación con el cliente, de las competencias, de las personas y, por supuesto, de la reorganización de los procesos. El objetivo es conseguir mayores beneficios aprovechando las ventajas y oportunidades que ofrecen las nuevas tecnologías digitales, para dotarse de procesos, estrategias y productos adecuados a un mercado altamente competitivo y rápido, donde, además, el cliente reclama atención personalizada.

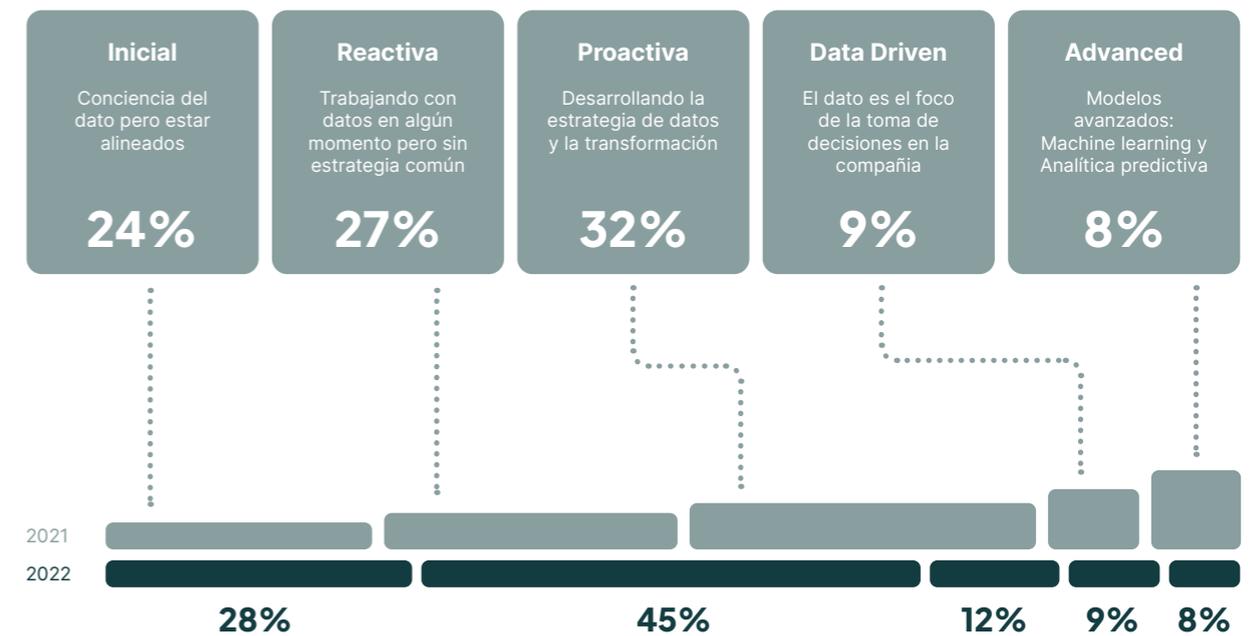
Desde hace ya algún tiempo, las empresas españolas están concienciadas de la necesidad de esta transformación, sin embargo, solo el 25% de ellas –en el caso de pymes es inferior– planifican desarrollar estrategias concretas y, de estas, solo el 50% tienen un presupuesto específico para ello. Por otro lado, aquellas empresas que se deciden y se ponen en marcha, encuentran grandes dificultades para cristalizar proyectos específicos que las hagan realmente avanzar.

La madurez data-driven de la empresa española

Algunas de las principales conclusiones de la encuesta «Madurez Data Driven» de la empresa española realizada por la empresa Incipy en el 2022 a 152 empresas españolas son:

- Un 97% de las empresas cree que la gestión con datos es una clara área de mejora en la empresa
- Solo el 36% de las compañías indican que la cultura del dato está desplegada en toda la organización. En ellas, las áreas de Negocio (42%), Finanzas (41%) y Marketing (40%) son las que gestionan y usan el mayor número de datos
- Un 25% de las empresas en España, todavía no poseen ninguna infraestructura de Business Intelligence
- Solo el 17% de las empresas son Data Driven y de ellas solo el 8% aplican modelos de analítica avanzada. El 32% de las empresas encuestadas indican que están ya desarrollando una estrategia de datos y la transformación cultural

Figura 11. Estadios de Madurez Data Driven



Por tanto, las empresas tienen claro que se ha de apostar por el dato y por tener equipos altamente cualificados en esta área, si bien, en líneas generales, el 50% de ellas están aún en los pasos iniciales del camino a ser consideradas empresas data-driven y solo un 17% hace un uso avanzado de los datos.

En resumen, el grueso de las organizaciones piensa que los datos deben estar en el centro de sus decisiones de negocio, pero la mayoría todavía usa la analítica de datos de forma aislada para abordar problemas de negocio específicos.

Los problemas actuales más habituales son: existencia de silos de información, falta de trazabilidad del dato entre los distintos aplicativos y herramientas de gestión, una infraestructura inadecuada de servidores, gestión ineficaz de datos maestros, creciente diversidad de datos, baja calidad e inexactitud del dato, falta de automatización de procesos susceptibles de serlo, falta de disponibilidad del dato necesario e históricos insuficientes para poder generar modelos, etc.

La Universidad data-driven

En este entorno, la universidad española en su conjunto, y cada una de las universidades que lo componen, son una organización que ha de afrontar esta misma evolución para poder seguir siendo uno de los ejes vertebradores de la sociedad.

No se trata de elegir entre enseñanza virtual o presencial, sino de potenciar ambas posibilidades, conviviendo de forma armonizada según necesidad, tal que den lugar a un aumento de la calidad de la enseñanza, apoyándonos en la oportunidad que nos ofrece el uso de los datos, la tecnología y el entorno digital en su conjunto.

Siendo así, ¿en qué punto del camino de madurez a ser organizaciones digitalizadas y data-driven nos encontramos? Además de compartir los problemas señalados para el resto de organizaciones y empresas, ¿qué particularidades debemos abordar desde la Universidad?

La mirada debe ser amplia, analizando el ciclo completo de procesos de forma transversal, desde la gestión administrativa hasta la propia docencia y los modelos de aprendizaje.

Digitalización de los procesos de gestión administrativa

La automatización de los procesos administrativos es parte fundamental de la digitalización que las universidades deben abordar. Se trata de generar la posibilidad de la experiencia digital completa en todas las fases de la vida universitaria, tanto para el alumno, como para el profesor, el personal administrativo y la institución en su conjunto.

Esto permitirá, en primer lugar, que la experiencia del estudiante sea más acorde a lo que el nativo digital espera, es decir, una comunicación virtual para hacer la mayoría de los trámites, agilidad en encontrar la información que necesita, agilidad en la interacción, agilidad en las respuestas... escenario necesario para fidelizar a nuestro «cliente».

Paralelamente, facilita el trabajo administrativo, al automatizarse los procesos de gestión de cobros, emisión de mensajes, de recordatorios, seguimiento de deudores, seguimiento de procesos de matriculación, convalidaciones, becas...; todo lo cual permite la reducción de los tiempos de gestión, reducción de costes e incremento de la productividad.

En añadido, la información generada y almacenada podrá ser utilizada para realizar análisis de situación y de seguimiento, así como para apoyar la toma de decisiones y generación de nuevas estrategias administrativas; todo ello con el objetivo de propiciar una mayor eficiencia administrativa.

Digitalización de los procesos de docencia y aprendizaje

Durante la pandemia de la Covid-19, digitalizar el aprendizaje se convirtió en una necesidad que nos obligó a reinventarnos, tratando de adaptarnos a una situación para la que solo algunos pocos estaban realmente preparados.

Si bien es cierto que llevamos años explorando e incorporando nuevas herramientas, tras la pandemia, podemos afirmar que la digitalización de la docencia es un imperativo; ha venido para quedarse.

En este sentido, se hace imprescindible contar con una plataforma educativa digital que ejerza de campus virtual, donde el alumno pueda acceder a sus estudios y a cada una de las asignaturas, así como a su contenido y materiales asociados. Se tratará de aportar una experiencia personalizada dando acceso a contenido añadido según intereses, permitiendo una comunicación ágil con el profesor y con el resto de los alumnos, permitiendo la creación de grupos de trabajo, haciendo llegar mensajes sobre eventos importantes para el alumno, etc.

Asimismo, el campus debe estar dotado de herramientas que faciliten las tareas del profesor. Facilitar la docencia virtual, mejorar la planificación de clases, exámenes y trabajos; organización de asignaturas y de alumnos por clases, creación de horarios; fácil seguimiento del rendimiento académico de cada alumno, al monitorear el cumplimiento de los trabajos propuestos, y permitir una evaluación más ágil y consistente; analizar la disponibilidad y el desempeño de los docentes; realización de encuestas de calidad, etc.

También aquí los datos así generados podrán retroalimentar el proceso, logrando enriquecer y mejorar el proceso en sí mismo, además de dar apoyo a la toma de decisiones estratégicas para la Universidad.



Personalizar la experiencia del aprendizaje

Incorporando la tecnología, plataformas y herramientas que permitan la digitalización de los procesos administrativos y de docencia, tendremos toda la información que necesitamos para conocer los patrones de comportamiento de nuestros estudiantes y así poder tomar decisiones que mejoren su experiencia.

Por ello, en primer lugar, podremos facilitar su aprendizaje tratando de adaptarnos a sus necesidades individuales mediante el empleo de distintas metodologías y recursos de aprendizaje, como contenidos visuales y textuales, clases invertidas, gamificaciones personalizadas, clases de apoyo, seguimiento de tutorías, diferentes ritmos de estudio... logrando así un mayor desarrollo del alumno a nivel personal y grupal, y contribuyendo a garantizar su éxito.

En segundo lugar, propondremos la generación de itinerarios formativos personalizados, donde el alumno elegirá el contenido curricular en base a sus características y circunstancias personales, sus intereses, motivaciones, objetivos profesionales y a la situación y demanda en el mercado.

De este modo lograremos maximizar la personalización de su formación y aprendizaje, tal que el estudiante sea parte activa de su educación, generando un aumento de su motivación y contribuyendo a la calidad y satisfacción percibida.

El Centro de Datos y de Analítica de la Universidad

Hemos hablado de cómo digitalizar todos los procesos de las distintas fases que componen la experiencia universitaria. Los datos así generados se han de almacenar y gestionar convenientemente, de forma centralizada, conectando e integrando las distintas plataformas y aplicativos de la Universidad, para poder ser posteriormente tratados y procesados para suministrar la información requerida a cada departamento. Esto requiere que apostemos por el despliegue de una arquitectura tecnológica híbrida centralizada que soporte todos los procesos que consideremos necesarios para poder llevarlo a cabo.

El objetivo es convertir la Universidad en una organización data-driven, y para ello, en primer lugar, todos los usuarios finales de las distintas áreas deben de tener acceso a la información que necesiten, en el modo que precisen y con la calidad adecuada, para que puedan tomar decisiones basadas en esos datos. Para conseguirlo, debemos desplegar cuatro ejes: gestión del dato, data governance, informacional y área de analítica avanzada.

El eje de gestión del dato o data management, la cual consiste en establecer los métodos y herramientas necesarias para recopilar, gestionar y publicar los datos que generamos, garantizando su trazabilidad, consistencia.

En segundo lugar, el área del Gobierno & calidad del dato, donde se desarrollan el conjunto de principios, procesos, funciones y responsabilidades que garanticen la consistencia de los datos, sobre los que se sustentan las decisiones empresariales, atendiendo a los atributos de calidad que debe cumplir cada dato.

El eje informacional, es aquel con el que aspiramos a medir toda la actividad de la universidad de forma transversal desde la capa administrativa hasta la docencia, y a su vez, logrando una visión 360 del alumno. La flexibilidad en el acceso y manejo de este entorno es necesario para facilitar el uso a todos los usuarios de la Universidad según su perfil de acceso y necesidades.

Y por último el eje de analítica avanzada, el cual nos va a permitir obtener un mayor beneficio de los datos a partir de la realización de modelos u algoritmos que nos permitan ofrecer un plus de servicios a los distintos departamentos y áreas, así como incidir positivamente en la estrategia de la Universidad.

Algunos ejemplos, de los nombrados en los apartados previos, de modelos que pueden llegar a realizarse en esta área, podrían ser:

- Analizar el comportamiento del estudiante en el campus virtual para apoyar su rendimiento académico y mejorar su experiencia, desarrollando modelos de clasificación basados en el análisis de los patrones de comportamiento y actividades diversas del alumno en la plataforma. Podremos maximizar el éxito de cada estudiante, así como detectar comportamientos fuera de la «normalidad»
- Prevenir el abandono mediante la creación de alertas tempranas que nos indican la probabilidad y causas de abandono de un estudiante. Supone la realización de un análisis predictivo con el que identificar las causas y corregir la situación mediante indicadores que reciben los profesores y tutores en el seguimiento del estudiante
- Definición de itinerarios o rutas de estudio personalizadas que se presentan como distintas opciones que ofrecen la mejor combinación entre intereses, habilidades del estudiante y la demanda del mercado laboral, maximizando su futura empleabilidad
- Predecir el grado de un éxito de un estudiante en su futura empleabilidad basándonos en las fortalezas del estudiante, sus logros, sus intereses profesionales y situación del mercado. Este indicador será de gran utilidad para profesores y tutores para adaptar la atención a cada alumno en función de las necesidades detectadas
- Desarrollo de servicios de atención al estudiante mediante chatbot para facilitar la interacción con el estudiante y ofrecerle un servicio de atención más personalizado, disponible las 24 horas. Respuestas a preguntas relacionadas con horarios, lugar de clase, notas, trabajos pendientes, instalaciones, biblioteca....
- Desarrollo de un conjunto de herramientas destinadas a mejorar el conocimiento y empoderamiento del profesorado. Incluye algoritmos para automatizar la corrección y calificación de tareas, panel de alertas por estudiante y profesor virtual que combina IA y software analítico

Para alcanzar este objetivo, además de arquitectura y procesos, necesitamos personas y equipos de trabajo. Los perfiles clave a incorporar en el proyecto son administradores de la base de datos, donde incluimos arquitectos, DataBase Administrator (DBAs), Data Quality y Data Governance, que se dediquen a asegurar la exactitud e integridad del dato entre los sistemas del dato y la base de datos analítica. Los ingenieros de datos son en su mayor parte ingenieros de bases de datos relacionales, pero actualmente también hay mucha información en entornos cloud, con lo cual hemos de tener en cuenta incluir a ingenieros de este entorno.

En la capa informacional sobre todo trabajamos con analistas de datos y de negocio, en el que se produce una combinación de conocimientos de estos dos perfiles, para poder generar, ya no el dato, sino información y conocimiento para las áreas usuarios finales.

Y en la parte de analítica avanzada, simplificando mucho, hablaremos del científico de datos, para poder construir modelos avanzados descriptivos, predictivos y prescriptivos, que sean modelos que puedan impactar en la eficiencia de la organización, tanto en los procesos como en los objetivos más estratégicos de la Universidad.

Lo que pretenderemos es tener entre todos, no como personas individuales, sino como equipo, esa figura con visión multidisciplinar de diferentes skills, tanto hard como soft skills, que debe de tener el equipo en sí, recreando esa figura que denominamos unicornio.

Uno de los puntos importantes que se nos plantea en las universidades es el reskilling de los técnicos de los actuales departamentos de informática. En este aspecto hemos de tener en cuenta que es viable sobre todo en el caso de los perfiles asociados a la gestión de bases de datos relaciones, Structured Query Language (SQL), arquitectos y DBA.

Para el resto de los perfiles comentados, que son más recientes en el mercado y se han formado ya desde el inicio en el manejo de software y lenguajes propios de estas áreas y del entorno digital, o bien perfiles como los data scientists que necesitan conocimiento de estadística y matemática, de inicio se antoja algo más difícil dicho reskilling, aunque no imposible. La clave está en que se trate de personas con una alta motivación intrínseca en desarrollarse en estos campos. La Universidad por su parte debe de tratar de identificar este talento, pues estos son de alto valor para la organización al conocer arquitecturas, procesos, organización, cultura...

Si optamos por la contratación externa, hemos de tener en cuenta que el mercado actual es un entorno altamente volátil y con alta inflación, es decir, hay escasez de oferta vs demanda, incluso para los perfiles más jóvenes, siendo además perfiles que les cuesta ser fieles a una organización pues sienten que «si están mucho tiempo en un mismo sitio se están perdiendo algo».

La Universidad debe resultarles un entorno atractivo para desarrollarse profesionalmente y para ello se han de desarrollar estrategias de fidelización, donde se contemple la formación continua y la actualización de sus conocimientos a las últimas novedades en el mercado.

Tras siglos de una aparente continuidad en la forma de educar en la Universidad, vivimos un momento disruptivo, que está propiciando una profunda transformación del concepto de aprendizaje y de la docencia en sí misma. Esto conlleva una actualización de la forma de operar en todas las áreas que abarca la Universidad, que ha de realizarse apoyándonos en las posibilidades que la tecnología y los datos aportan en la sociedad digital actual.

La Universidad no ha de sentir invulnerabilidad ante este nuevo paradigma. Aquellas organizaciones que sean más ágiles en promover y adoptar este cambio serán las que liderarán la creación del nuevo paradigma educativo, que bien pudieran ser universidades privadas, grandes corporaciones u otro tipo de instituciones que realicen las inversiones adecuadas. Es el momento de actuar y de liderar la nueva realidad.



COORDINADOR

Juan Camarillo Casado

Universidad de Sevilla

EQUIPO EDITORIAL

Grupo de trabajo de directores
TI de Crue-Digitalización

Lluís Alfons Ariño Martín

Coordinador TIC de la gerencia y
director del Servicio de Informática
de la Universitat Rovira i Virgili

Adelaida Cabrero Bueno

Jefa del Servicio de Informática
de la Universidad de Jaén

Joaquín Canca Cuenca

Director técnico del Servicio de
informática de la Universidad de
Málaga

Andrés J. Prado Domínguez

Director del Área de Tecnología y
Comunicaciones de la Universidad
de Castilla-La Mancha

AUTORES

Joaquín Canca Cuenca

Director técnico del Servicio de
informática de la Universidad de
Málaga

Carlos Alonso Peña

Director de la División Oficina del Dato,
Secretaría de Estado de Digitalización e
Inteligencia Artificial

Lluís Alfons Ariño Martín

Coordinador TIC de la gerencia y director
del Servicio de Informática de la Universitat
Rovira i Virgili y miembro del Advisory Board
de DS4Skills

Ricard Martínez Martínez

Director de la Cátedra de Privacidad
y Transformación Digital Microsoft-
Universitat de València

Ismael Caballero

Profesor titular de la Universidad de
Castilla-La Mancha

Nuria Canal

Chief Data Officer Member



 **crue** Universidades
Españolas Digitalización