



Alfabetización en datos: percepción y grados de satisfacción de estudiantes universitarios

Yolanda Martín González¹; Crispulo Travieso Rodríguez²; José C. Toro Pascua³; Ana Iglesias Rodríguez⁴; Azucena Hernández Martín⁵

Recibido: 20 de noviembre de 2022/ Aceptado: 15 de octubre de 2023

Resumen. Este artículo presenta los resultados obtenidos de un estudio realizado en el marco de un Proyecto de Innovación Docente sobre alfabetización en datos, desarrollado en cinco titulaciones de grado de la Universidad de Salamanca, con el objeto de conocer la percepción y el grado de satisfacción de los estudiantes universitarios respecto a la búsqueda y recuperación de datos abiertos. Se ha aplicado el enfoque metodológico cuantitativo no experimental, empleando técnicas estadísticas, descriptivas e inferenciales. Se ha hecho una selección muestral por disponibilidad accediendo a 310 discentes de las ramas de conocimiento Ciencias Sociales y Jurídicas, y Artes y Humanidades. El instrumento de recogida de datos ha sido un cuestionario conformado por preguntas de respuesta cerrada y abierta, referidas a las variables que se pretendía estudiar. Los resultados obtenidos han permitido identificar las dificultades que encuentran los estudiantes en cuanto a la elección de la fuente de datos, la ejecución del proceso de búsqueda de información y la comprensión de los dataset. Asimismo, ponen de manifiesto cómo los discentes valoran positivamente la formación en datos y su utilidad en el aprendizaje de otras asignaturas, así como en su vida profesional. El estudio efectuado ha obtenido evidencia empírica acerca del nivel de percepción y satisfacción que poseen los estudiantes de grado en cuanto a la alfabetización en datos. Además, ha expuesto la necesidad de incorporar esta competencia transversal en la formación universitaria y de promover la, hasta ahora, limitada investigación sobre el uso potencial que tienen los datos abiertos en el aprendizaje.

Palabras clave: Alfabetización en datos; Educación superior; Alfabetización informacional; Estudiante universitario de primer ciclo; Competencias digitales.

[en] Data Literacy: Perception and degree of satisfaction among university students

¹ Universidad de Salamanca. Departamento de Biblioteconomía y Documentación
E-mail: ymargon@usal.es

² Universidad de Salamanca. Departamento de Biblioteconomía y Documentación
E-mail: ctravieso@usal.es

³ Universidad de Salamanca. Departamento de Biblioteconomía y Documentación
E-mail: jctoro@usal.es

⁴ Universidad de Salamanca. Departamento de Didáctica, Organización y Métodos de Investigación
E-mail: anaiglesias@usal.es

⁵ Universidad de Salamanca. Departamento de Didáctica, Organización y Métodos de Investigación
E-mail: azuher@usal.es

Abstract. This article presents the results obtained from a study carried out within the framework of a Teaching Innovation Project on data literacy, developed in five undergraduate degrees at the University of Salamanca, with the aim of finding out the perception and degree of satisfaction of university students with respect to the search and retrieval of open data. The non-experimental quantitative methodological approach was applied, using statistical, descriptive and inferential techniques. A sample selection has been made by availability, accessing 310 students from the branches of knowledge Social and Legal Sciences, and Arts and Humanities. The data collection instrument was a questionnaire made up of closed and open-ended questions, referring to the variables to be studied. The results obtained made it possible to identify the difficulties encountered by the students in terms of choosing the source of data, carrying out the information search process and understanding the datasets. They also show how the students value positively the training in data and its usefulness in the learning of other subjects, as well as in their professional life. The study has obtained empirical evidence about the level of perception and satisfaction that undergraduate students have regarding data literacy. It has also exposed the need to incorporate this transversal competence in university education and to promote the, so far, limited research on the potential use of open data in learning.

Keywords: Data literacy; Higher education; Information literacy; Undergraduate student; Digital skills.

Sumario. 1. Introducción. 2. Método. 3. Resultados y discusión. 4. Conclusiones. 5. Referencias bibliográficas.

Cómo citar: Martín González, Y. [et al.] (2023). Alfabetización en datos: percepción y grado de satisfacción de estudiantes universitarios, en *Revista General de Información y Documentación* 33 (2), 353-371.

1. Introducción

La situación de alerta sanitaria provocada por la COVID-19 ha puesto de relieve la necesidad de contar con las competencias digitales suficientes para desenvolvemos en una sociedad virtual. Entre esas competencias y habilidades requeridas en el contexto digital actual, se incluye la alfabetización informacional, que pretende desarrollar la capacidad de localizar y recuperar datos e información fiable y de calidad en cualquiera de sus formas (textual, gráfica, multimedia, etc.).

En el Marco de Competencia Digital para los ciudadanos (Comisión Europea, 2017), la primera área competencial es, precisamente, Información y alfabetización digital. Con ello se indica la imbricación existente entre las competencias digitales y la alfabetización informacional (ALFIN), donde una persona ha de ser capaz de identificar la necesidad de información; navegar, buscar y filtrar información, datos y contenidos digitales; evaluar de manera crítica la información; y organizar, almacenar y recuperar la información, los datos y los contenidos digitales (Carretero et al., 2017).

ALFIN comprende todas las habilidades y aptitudes que posibilitan a una persona resolver de forma eficaz y ética sus necesidades de información, con independencia del formato en el que se encuentren los datos, contenidos y fuentes que requiere para ello, así como evaluar y gestionar la información. Dicha información se puede entender como “(...) el conjunto de datos organizados y contextualizados; que se transmiten en el proceso de comunicación social; y que una vez asimilados permiten la solución de problemas, la toma de decisiones y la disminución de incertidumbre

(Pozo-Jara, 2017: 4). Atendiendo a esta definición podríamos referir de forma literal que la alfabetización informacional se relaciona con las habilidades básicas para acceder a la información en diferentes formatos y con utilidad práctica para las personas (Díaz-Arce y Loyola-Illescas, 2021).

1.1. Alfabetización de datos

Para explicar qué es la alfabetización en datos resulta necesario definir previamente el objeto de dicha alfabetización, el dato, ya que son pocas las definiciones que engloban este aspecto que, en nuestra opinión, es tan relevante. Si no delimitamos el concepto dato podemos encontrarnos con problemas a la hora de diseñar la actividad formativa y determinar, por tanto, el propósito de la misma.

Un dato puede definirse como la unidad informativa más pequeña, que per se no posee ningún significado, que se ofrece en bruto y que es tratable de manera automatizada. Pero, al mismo tiempo, esta definición podría modificarse y ampliarse más allá de la capacidad de predicción que proporciona el aprendizaje automático desde un punto de vista cuantitativo, si se tuvieran en cuenta sus características básicas, su recopilación, almacenamiento y protección, momento en el que los datos pasarían a concebirse como algo más que una cosa o un objeto, ya que, por un lado, permitirían explicar y explorar, desde un punto de vista cualitativo, su singularidad y las anomalías detectadas que no pueden explicarse con números (datos); y, por otro lado, posibilitarían interpretar y dar sentido a los patrones inesperados que podemos encontrar en el análisis de los Big Data (Markham, 2017).

Dada la dificultad que entraña definir el término dato, quizás se debería valorar, desde el punto de vista cualitativo, qué es lo que cada persona o qué es lo que los ciudadanos entienden e interpretan que son los datos y, en consecuencia, deciden qué es lo que cuenta como datos, puesto que la conceptualización que realicen sobre el mismo va a depender de la experiencia que cada uno de ellos haya vivido con respecto a, a través de o dentro de los datos.

A este respecto, Markham (2019) dice que:

This can be achieved by helping people become qualitative researchers of their own experience, to identify how identities, events, and cultural formations are quantified, and more importantly, to interrogate the ways their personal behaviors and activities and relationships—in the form of information—might be harvested, stored, and eventually commoditized for sale in the corporate or governmental marketplace. (p.754)

Para lograr este cometido, desde el punto de vista educativo, sería importante formar a tantas personas como sea posible para que puedan hacer un uso crítico y adecuado de la ingente cantidad de datos digitales que invaden nuestras vidas, así como de la manipulación, recopilación y comunicación que se hace de los mismos.

Hemos de tener presente que en el ámbito educativo los datos son, como afirman Bradbury y Roberts-Holmes (2017), una “fuerza productiva”, que en el caso de la alfabetización, prioriza las relaciones de causa-efecto en lo que respecta a los logros (estadísticamente hablando), olvidándose de otras dimensiones importantes que tienen que ver con las experiencias de los docentes y de los estudiantes en materia

de alfabetización, ya que se centran más en indicadores del progreso limitados y medibles (Hayes et al., 2017).

Es importante, asimismo, destacar que los datos no son solo significativos en lo que respecta a lo que ocurre en los centros educativos, sino también la forma en que se visualizan y difunden (Grant, 2017).

Si bien es cierto que la visualización de los datos ayuda y facilita su interpretación, también es verdad que la significación, la abstracción y la descontextualización de los mismos (O'Halloran, 2008), da lugar a un desequilibrio de poder y a una lucha por el significado que privilegia a ciertas personas, instituciones y perspectivas sobre otras, ya que no hay nada que pueda ser considerado un objeto neutro que simplemente se pueda observar y describir, entrando en juego otras dimensiones y circunstancias que deberían contemplarse desde el punto de vista ético, dada la repercusión y fuerte impacto que estas visualizaciones podrían tener a nivel social, educativo y profesional en un futuro (Markham, 2019).

Como podemos observar, el concepto de Alfabetización en datos se encuentra directamente vinculado al concepto de alfabetización informacional, y alude a la necesidad de utilizar mejor los datos y de realizar descubrimientos en base a ellos. La data literacy se ha desarrollado, en especial, en los últimos años, gracias al impulso recibido de la nueva ciencia de datos, la analítica de datos, el big data y el open data.

La alfabetización en datos comprende la habilidad para recopilar, leer, argumentar, utilizar, recuperar, analizar, comprender, comunicar, visualizar, evaluar críticamente y reutilizar datos.

1.2. Alfabetización en datos en contextos universitarios

Si hacemos un análisis pormenorizado de los contextos universitarios en general, podemos observar cómo en los últimos diez años y, de manera especial, durante la situación sanitaria provocada por la COVID-19, se están produciendo en ellos cambios y ajustes significativos en los procesos de digitalización que implican un uso indiscriminado de datos (Daniel, 2014), que repercuten especialmente en los procesos académicos y de investigación. Así pues, la educación superior está cada vez más datificada, lo que no significa que los datos sean un elemento de reciente aparición en el contexto universitario, pero sí que el estado de los datos ha cambiado, se ha digitalizado, constituyendo en los últimos años, una característica definitoria de la educación superior (Selwyn, Henderson y Chao, 2015).

Este hecho pone de manifiesto la necesaria formación que, tanto estudiantes como docentes e investigadores, precisan en alfabetización en datos para evitar prácticas discriminatorias y de control que puedan repercutir negativamente en su trayectoria académica y profesional (Bradbury, 2018; Manolev, Sullivan, & Slee, 2018; Lewis, & Holloway, 2018; Raffaghelli, 2019; Autor/a & Autor/a, 2020; Autor/a & Autor/a, 2022).

Llevar a cabo una formación centrada en prácticas educativas basadas en evidencias (Gairín, & Ion, 2021), sin duda alguna, favorecerá, potenciará y

beneficiará la compartición de dinámicas y situaciones que se generan en la investigación y en la enseñanza, en cuanto a la manipulación y comunicación de datos, y el análisis, evaluación y compartición de prácticas pedagógicas efectivas, de los que podrían beneficiarse investigadores, docentes y estudiantes, utilizando los “[...] datos educativos como contenido [...] para el avance de la enseñanza como parte de la ciencia educativa” (Raffaghelli, 2019: 181).

En el ámbito académico, los programas ofertados en materia de alfabetización en datos tienen como objetivo el conocimiento y uso de los datos como recurso para la investigación científica; sin embargo, en el contexto de las organizaciones, se ofrece formación que facilite la adquisición de competencias y habilidades necesarias para tomar decisiones informadas en base a datos (Hyatt, 2021). Es en el primer caso, donde se relaciona data literacy con information literacy mientras que, en el segundo, se enmarca en la analítica de datos y la ciencia de datos.

Los programas formativos ofrecen, por lo general, los siguientes elementos: 1) definición de alfabetización en dato; 2) competencias que comprende; 3) conocimientos, destrezas o habilidades integradas en el programa.

La acción formativa en alfabetización en datos contemplada en el Proyecto de Innovación Docente, Integración de la alfabetización en datos en el programa curricular de asignaturas de grado mediante la metodología Aprendizaje basado en datos, desarrollado en la Universidad de Salamanca y en el que se enmarca el presente estudio, se encuentra en la línea marcada por la Conferencia de Rectores de Universidades Españolas (CRUE) para la implementación de la Agenda 2030 y el cumplimiento de los Objetivos de Desarrollo Sostenible (CRUE, 2020). En concreto, la instrucción ofrecida a los estudiantes nos ha permitido integrar el ODS 4 Una educación de calidad que favorece la educación inclusiva y en competencias transversales para una ciudadanía global.

2. Método

Puesto que nuestro propósito era analizar el estado de un hecho en un contexto determinado, para esta investigación se optó por un enfoque metodológico cuantitativo no experimental, empleando técnicas estadísticas, tanto de tipo descriptivo como inferencial. A continuación, se detallan las distintas fases de la secuencia metodológica seguida.

2.1. Objetivos

El principal objetivo de este estudio es analizar la percepción y el grado de satisfacción de los estudiantes de grado universitario respecto a la acción formativa recibida en torno a la alfabetización en datos.

Partiendo de este objetivo general, el estudio pretende alcanzar, asimismo, los siguientes objetivos específicos:

- Identificar las dificultades que encuentran los estudiantes en cuanto al manejo de los datos abiertos.

- Saber qué utilidad creen los discentes que pueden tener los datos en el aprendizaje de otras asignaturas.
- Conocer la experiencia previa que los estudiantes tienen en cuanto a la búsqueda y recuperación de datos abiertos.
- Averiguar de qué modo piensan los estudiantes de grado que pueden emplear en un futuro las fuentes y técnicas de alfabetización de datos.

2.2. Población y muestra

Como se ha mencionado anteriormente, este estudio se llevó a cabo en el marco de un Proyecto de Innovación Docente multidisciplinar. Esta posibilidad de contar con una muestra perteneciente a diversos ámbitos de conocimiento era uno de los elementos que se quería poner en valor expresamente, pues se partía de la premisa de la utilidad transversal de esta materia en la formación universitaria. Por ello, se realizó una selección muestral por disponibilidad, accediendo a estudiantes pertenecientes a diferentes grados dentro de las ramas de Ciencias Sociales Jurídicas, y Arte y Humanidades, de los que todos los autores de este estudio eran docentes. En concreto, se eligieron cinco Titulaciones: Grado en Criminología (G-CRI), Grado en Educación Social (G.ES), Grado en Información y Documentación (G-ID), Grado en Lenguas, Literaturas y Culturas Románicas (G-LLCR) y Grado en Pedagogía (G-P). El total de sujetos participantes, según el muestreo no probabilístico, fue de 310 y su distribución por Titulaciones se muestra en la Tabla 1.

Tabla 1. Participantes según titulación.

Titulación	Nº Participantes	Porcentaje
Grado en Criminología	165	53.2%
Grado en Educación Social	3	1%
Grado en Información y Documentación	18	5.8%
Grado en Lenguas, Literaturas y Culturas Románicas	14	4.5%
Grado en Pedagogía	110	35.5%

La composición de la muestra participante estaba en parte determinada por el número de matriculados en los correspondientes grados; fue mayoritario el grupo de Criminología, con más de la mitad de la población estudiada, seguida de la titulación en Pedagogía, con un 35.5%.

En lo referido al nivel cursado, la mayor participación se dio en los primeros cursos, a los que pertenecían casi dos tercios de los alumnos de este estudio (a primero y segundo curso correspondían el 20.6% y el 45.2% del total, respectivamente), mientras que a los últimos niveles pertenecían el 34.2% restante, distribuidos de forma homogénea entre tercer y cuarto curso, con 54 y 52 estudiantes, respectivamente.

Como elemento vinculado, pero no estrictamente determinado por el curso, también se recopiló la variable edad, estableciendo las franjas reflejadas en la Tabla 2.

Tabla 2. Participantes según edad.

Titulación	N.º Participantes	Porcentaje
18-19 años	147	47.4%
20-22 años	137	44.2%
Más de 22 años	26	8.4%

Como última variable sociodemográfica, los datos desglosados según el género de los participantes demostraron que las mujeres representaban una clara mayoría (83.5%) -situación esperable por su preeminencia en las aulas universitarias, especialmente en las ramas de conocimiento escogidas (Gómez Marcos, Vicente Galindo y Martín Rodero, 2019)-, frente al género masculino (15.5%) u otros géneros (1%).

Instrumento de recogida de datos

Para comprobar sus conocimientos previos sobre la alfabetización en datos, así como su reacción tras una propuesta de formación en esta materia, se optó por la elaboración de un cuestionario como instrumento de recogida de información. Su estructura estaba compuesta por varios datos sociodemográficos básicos (titulación, curso, edad y género), seguidos por diez preguntas relacionadas directamente con el objeto de estudio, en las que se combinaba la modalidad de respuesta cerrada y abierta.

Estas preguntas hacían referencia a las variables que se pretendía estudiar y que aludían a los siguientes aspectos sobre alfabetización en datos:

- Utilidad y dificultad de una tarea específica.
- Tipos de datos empleados para su resolución.
- Experiencia previa en tareas y contenidos similares.
- Proyección y empleo futuro de fuentes y técnicas de alfabetización de datos.

Antes de aplicar el cuestionario, y para proceder a su validación, se utilizó en primer lugar el juicio de expertos. Para ello se solicitó a tres docentes universitarios expertos tanto en el tema objeto de estudio como en metodología de investigación que valorasen el instrumento atendiendo a la coherencia, relevancia y claridad de cada ítem. En segundo lugar, se llevó a cabo una prueba piloto del cuestionario con un grupo previamente seleccionado de estudiantes de Criminología. Ambos procedimientos permitieron la depuración del cuestionario y contar con su versión definitiva.

Por último, cabe señalar que se han seguido los protocolos de bioética estipulados por la Universidad de Salamanca.

2.3. Procedimiento de administración del cuestionario

Debido al carácter novedoso del empleo de técnicas de alfabetización en datos, se consideró oportuno establecer varias fases en la administración del cuestionario. En primer lugar, cada uno de los docentes implicados planteaba una sesión formativa teórica de 3 horas de duración en la que se explicaban, en el marco de la asignatura, el concepto, origen y tipología de los datos; el uso informacional de los datos; su acceso en abierto; y su reutilización. A continuación, se les planteaba una actividad práctica basada en la metodología de resolución de problemas que consistían en hipotéticas necesidades de información que debían resolver mediante la búsqueda y recuperación de datos disponibles en repositorios digitales. La última fase consistía en la cumplimentación del cuestionario por parte de los estudiantes, proceso que se realizó en formato electrónico una vez terminada la actividad.

Las asignaturas participantes en el estudio fueron las siguientes: Big Data en Criminología, Comunicación en Educación, Fuentes Especializadas de Información, Modelos y estrategias de formación del profesorado, Recursos documentales e informáticos, y Seminario de Orientación e Integración Laboral II.

En cuanto a la programación temporal, esta vino determinada por la propia calendarización y acceso de las materias de cada titulación incluida en el estudio, dado que se pretendió integrar esta actividad dentro del programa formativo, y que no figurara como una actividad específica percibida como adicional por los participantes. Ello explica que el periodo de recogida de datos se extendiera entre los meses de enero y mayo de 2021.

2.4. Análisis y tratamiento de los datos

Para la distribución del cuestionario se empleó la aplicación Google Forms, siendo posteriormente los datos volcados para su tratamiento en los programas Excel y en el paquete estadístico SPSS versión 25. A continuación, se llevaron a cabo los análisis descriptivos de las respuestas dadas a los diferentes ítems y se exploró la relación existente entre las variables género, curso y titulación y el resto de las cuestiones.

3. Resultados y discusión

Teniendo en cuenta que el objetivo primordial era comprobar el grado de satisfacción de los participantes hacia la actividad formativa, se ha procedido a realizar un estudio descriptivo de los resultados de cada ítem del cuestionario, además de analizar, mediante estudios inferenciales, los posibles factores o variables de identificación, como la titulación, el curso y el género que pueden condicionar las respuestas dadas a los diferentes ítems de dicho cuestionario. Para comprobar el grado o no de significación estadística, se efectuó la prueba de chi cuadrado (X^2), teniendo en cuenta las características de las variables consideradas.

Se han presentado los distintos resultados organizándolos en torno a los cuatro aspectos sobre alfabetización en datos ya indicados en la metodología.

3.1. Utilidad y dificultad de una tarea específica

El 92% de los estudiantes encuestados consideraron que la tarea realizada les había resultado útil. Fundamentalmente, en un 56% de los casos, porque aprender sobre el uso y recuperación de datos puede ser beneficioso para su vida profesional. El 39.8% estimaron que podrían utilizar el conocimiento adquirido para su vida académica, mientras que el 4.2% estiman más el empleo en su vida personal.

Atendiendo a la titulación universitaria se apreciaron diferencias estadísticamente significativas en seis de los diez ítems del cuestionario: 1 y 1b, 2 y 2b, 4, 5, 7 y 10. Solo presentaremos en este apartado las relativas al contenido del mismo.

Tabla 3. Participantes de la actividad realizada y beneficios.

	G-CRI (N= 165)	G-ES (N=3)	G-ID (N=18)	G-LLCR (N=13)	G-P (N=111)	X ²	p
Sí	56.4%	1.1%	4.3%	3.5%	34.8%		
No	19.2%	-	23.1%	11.5%	46.2%	25.80	.001
Beneficios en la vida profesional	56%	1.3%	7.7%	3.1%	35.2%		
						58.15	.000
Beneficios en la vida académica	60%	-	3.5%	1.8%	34.5%		
Beneficios en la vida personal	33.4%	-	8.3%	-	25.22		

Fueron los estudiantes de Criminología quienes consideraron más útil la actividad efectuada (Ítem 1). Los principales beneficios que adujeron (Ítem 1b) tienen que ver con su vida profesional y académica (titulaciones de Criminología, Lenguas, Literaturas y Culturas Románicas y Pedagogía), y con aspectos profesionales y personales (Información y Documentación).

Considerando el curso en el que se encontraban, las diferencias significativas se situaron en los ítems 1 y 1b, 2b, 5 y 5c, y 10. En la tabla 4 se observan los principales resultados, relacionados con la utilidad de la tarea efectuada y sus beneficios.

Tabla 4. Utilidad de la actividad realizada y beneficios.

	Primer curso (N=63)	Segundo curso (N=140)	Tercer curso (N=54)	Cuarto curso (N=51)	X ²	p
Sí	17.7%	46.1%	19.1%	17%	23.78	.002
No	53.8%	34.6%	0	11.1%		
Beneficios en la vida profesional	3%	45.9%	19.5%	16.4%	26.31	.010
Beneficios en la vida académica	17.7%	47.8%	18.6%	15.9%		
Beneficios en la vida personal	16.7%	33.3%	16.7%	33.3%		

Los alumnos de segundo (46.1%), tercero (19.1%) y cuarto cursos (17%) fueron los que más utilidad atribuyeron a la actividad formativa realizada, frente a los de primero con un porcentaje importante de respuestas negativas a este primer ítem (53.8%) debido, probablemente, a que todavía no han tenido una trayectoria académica suficiente como para intuir o percibir más claramente que sus compañeros de cursos superiores, el provecho que puede tener la actividad efectuada. Ello explicaría también que piensen que los beneficios de la misma se puedan producir más en el aspecto académico (17.7% de primero y 47.8% de segundo), frente a los que cursan, por ejemplo, cuarto que aluden, además, a otros de tipo personal (33.3%) y profesional (16.4%).

Otra variable considerada, como ya se indicó, ha sido el género. Como se aprecia en la tabla 5, solo se han identificado diferencias en el ítem 1. Así, las mujeres atribuyeron mayor beneficio a la actividad (4.4%) que los hombres (15.2%) y otros colectivos (0.4%). Aunque, como ya se ha señalado, la muestra de mujeres es mayor que la de hombres y la categoría otros, la diferencia evidenciada por chi cuadrado ha de entenderse en términos de género, al tener en cuenta en el cálculo del estadístico el tamaño de la muestra.

Tabla 5. Utilidad de la actividad realizada.

Ítems		Hombre	Mujer	Otros	Chi cuadrado	p
Utilidad de la actividad realizada	Sí	15.2%	84.4%	0.4%	14.545	.002
	No	19.2%	73.1%	7.7%		

A pesar de la opinión generalizada de la utilidad de la actividad, al 77% de los encuestados les resultó difícil localizar los datos para resolver los ejercicios

propuestos (Ítem 2). Fundamentalmente -en un 56% de los casos- porque desconocían el proceso de búsqueda dentro de la fuente o recurso de información. Además, un 27% encontraron la dificultad en la selección de la fuente de información para resolver las prácticas. Esto nos lleva a pensar que mayoritariamente fueron capaces de localizar la herramienta adecuada para identificar los datos, aunque no tenían las competencias necesarias para localizar la información solicitada en los ejercicios dentro de cada repositorio de datos. Los estudiantes que afirmaban haber tenido menores dificultades, de acuerdo con el estudio inferencial realizado, fueron los alumnos de Criminología (58.6% de respuestas negativas), además de los de Lenguas, Literatura y Culturas Románicas (15.7%). Los momentos en los que encontraron esta dificultad (Ítem 2b) fueron en la comprensión del dataset y durante el proceso de búsqueda en el contexto de la fuente de información, por parte de las Titulaciones de Criminología (46.3% y 43.7%) y Pedagogía (51.2% y 44.4%); mientras que, en el G-ID y el G-LLCR, se observó un mayor problema en la selección de la fuente o recurso de información. Se ha de reseñar también que, atendiendo a los porcentajes que se incluyen en la tabla 5, el alumnado de Pedagogía fue el que tuvo especialmente más problemas en todos los momentos de la tarea. Ello puede deberse a que no han recibido ningún tipo de instrucción específica sobre búsqueda de información en recursos digitales, a excepción de la acción formativa realizada con anterioridad a la cumplimentación del cuestionario. Los estudiantes de Criminología y de Información y Documentación, aunque también han manifestado dificultades, estas fueron menores, comparativamente, en porcentajes de respuesta, porque poseen una formación previa en alfabetización informacional. Por tanto, se hallan familiarizados con la metodología de búsqueda en repositorios, si bien no de datos abiertos.

Tabla 6. Dificultad para encontrar datos y momento.

	G-CRI (N=165)	G-ES (N=3)	G-ID (N=18)	G-LLCR (N=13)	G-P (N=111)	X ²	p
Sí	51.3%	1.3%	7.1%	0.8%	39.8%		
No	58.6%	-	1.4%	15.7%	24.3%	38.27	.000
Comprensión de Dataset	46.3%	-	2.4%	-	51.2%		
Proceso de Búsqueda dentro de fuente de información	43.7%	-	2.1%	1.4%	44.4%	58.69	.000
Selección de la fuente o recurso	7.1%	-	5.8%	2.9%	20.3%		

Las respuestas se diversifican bastante considerando el curso (Ítem 2b). Así, los estudiantes de primero se refieren a la comprensión de los dataset como la actividad

más compleja (41.5%) (Tabla 6). Los de segundo (45.1%) y tercero (26.8%) situaron la dificultad en el proceso de búsqueda dentro de la fuente de información. Y el alumnado de cuarto curso encontró más problemas en la selección de la fuente o recurso (24.6%). Se ha analizado el porcentaje de respuesta a esta cuestión otorgada por los alumnos de las titulaciones consideradas en cada uno de los cursos. El 51.2% de la muestra de Pedagogía corresponde al primer curso, lo que explicaría que tuvieran más dificultad a la hora de comprender los dataset porque no han trabajado previamente con este tipo de registros. Parece lógico pensar, como se ha señalado en otro momento, que a medida que se va avanzando en los diferentes cursos, los alumnos Pueden haber estudiado ya en algunas materias aspectos relacionados con la alfabetización en datos, y muestren, en consecuencia, dificultades más específicas.

Tabla 7. Dificultad para encontrar datos y momento.

	Primer curso (N=63)	Segundo curso (N=140)	Tercer curso (N=54)	Cuarto curso (N=51)	X ²	p
Comprensión de Dataset	41.5%	24.4%	12.2%	22%	58.88	.000
Proceso de búsqueda dentro de fuente de información	21.1%	45.1%	26.8%	7%		
Selección de la fuente o recurso	13%	52.2%	10.1%	24.6%		

3.2. Tipos de datos empleados para su resolución

Los estudiantes respondieron mayoritariamente, con un 63% que los datos obtenidos fueron numéricos en su totalidad. Sin embargo, de las 14 preguntas que debieron resolver, tan solo 5 se solucionaban con datos cuantitativos, mientras que las 9 restantes se resolvían con respuestas textuales. Agrupando y normalizando las opciones obtenidas, se comprueba cómo únicamente 9 estudiantes -un 3%-, afirmaron que los datos que obtuvieron fueron numéricos y textuales, que fueron de todo tipo o bien que dependía de la pregunta, opciones que eran las más correctas. Evidentemente puede tratarse de un problema de comprensión de la tipología de datos, o de falta de atención en el momento de responder la pregunta. Durante la ejecución de la práctica, ocho de las respuestas se obtenían con la consulta de ficheros CSV o XLS. Mientras que seis de las catorce preguntas se podían resolver con la consulta en diferentes páginas webs. En este caso, los estudiantes sí que respondieron la respuesta correcta de consulta de datos en formato XLS o XLSX,

con un 58.4%. Destacamos la opción de formato PDF, que fue manifestada incorrectamente por el 14.8% de los estudiantes, ya que ninguna cuestión del ejercicio se solucionaba con archivos de esta tipología.

Los estudiantes que emplearon formatos de consulta de datos (Ítem 4) más diversos fueron los correspondientes a las titulaciones de Criminología y de Pedagogía y, seguidamente, el alumnado de Lenguas, Literatura y Cultura Románicas y de Información y Documentación. Los participantes de Criminología utilizaron mayoritariamente el formato DOC (79.3%) y XSL (55%), este último, como se ha indicado, correcto para la resolución de la tarea. JPG, incorrecto para el desarrollo de la actividad, fue usado más en la actividad realizada por el alumnado de Pedagogía (46.7%), Lenguas, Literatura y Culturas Románicas (13.3%), así como en Educación Social (6.7%). El formato PDF lo empleó un porcentaje importante de estudiantes de los distintos grados a la hora de responder a las preguntas del cuestionario. Todo ello evidencia, en general, un insuficiente dominio de los formatos de consulta correctos para responder a las diferentes preguntas de la práctica realizada con posterioridad a la acción formativa. Tengamos en cuenta la elevada proporción de alumnado que empleó, por ejemplo, el formato .PDF, que no era necesario para la resolución de ninguna cuestión. O el formato JPG, especialmente destacado por los alumnos de Pedagogía, que tampoco era preciso emplear en la práctica. Estos resultados, junto a otros, permiten vislumbrar la necesidad de un mayor tiempo de formación, con el que el alumnado pueda ir interiorizando y asimilando toda la información teórica y práctica desarrollada durante la acción formativa. No podemos olvidar que el 92% consideró que la tarea les había resultado útil, por lo que no es una cuestión de falta de interés, sino de insuficiencia a nivel formativo.

Tabla 8. Formatos de consulta de datos.

	Criminología (N= 165)	Educ. Social (N=3)	Informac. y Documentación (N=18)	Lenguas, literatura y Románicas (N=13)	Pedagogía (N=111)	X²	p
DOC	79.3%	-	-	-	20.7%		
HTML	31.3%	-	12.5%	12.5%	43.3%		
JPG	33.3%	6.7%	-	13.3%	46.7%	97.49	.011
XSL	55%	-	6.5%	0.6%	38.9%		
PDF	42.2%	2.2%	15.6%	13.3%	26.6%		

Nota: Solo se han incorporado en la tabla los formatos en los que había algún porcentaje de respuesta

3.3. Experiencia previa en tareas y contenidos similares

Un 78% de los encuestados no habían utilizado los datos como recurso para el aprendizaje en ninguna asignatura. Tan solo 68 estudiantes afirmaban que sí se habían formado en recuperación de datos en otras asignaturas, fundamentalmente en Metodología de la Investigación en Ciencias Sociales, Recursos Documentales e Informáticos y en Teorías Criminológicas, todas ellas asignaturas del grado en Criminología, alumnado mayoritario en el estudio. El tipo de fuentes de recuperación de datos que declararon haber utilizado son de información estadística, en un 35.8%, y buscadores generales de datos, en un 33.9%. También, un 17% señalaron haber empleado repositorios de datos de investigación, y un 13% portales institucionales de datos abiertos. No se han encontrado diferencias estadísticamente significativas atendiendo a la titulación, el curso y el género en lo que respecta a la experiencia previa en tareas de este tipo y contenidos similares.

3.4. Proyección y empleo futuro de fuentes y técnicas de alfabetización de datos

El porcentaje de encuestados que creen que podrán utilizar las fuentes de recuperación de datos para sus trabajos académicos fue de un 73.6%, mientras que tan solo un 2.6% consideraron que no los utilizarán, un 80% de ellos, por la dificultad en encontrar los datos. El resto manifestaron que no sabían si emplearán estas fuentes para sus tareas académicas. Además, tras la ejecución de la actividad, un 95.8% opinaron que sería conveniente incorporar la utilización de los datos en más asignaturas. Los propios estudiantes son conscientes de la necesaria formación en torno a la alfabetización en datos, una necesidad que han evidenciado distintos especialistas, y que ya se puso de manifiesto en la parte teórica del presente trabajo.

Los participantes de Criminología (73.5%) e Información y Documentación (7.4%) afirmaron, en mayor grado, que la utilización de los datos puede ser un recurso para el aprendizaje de otra asignatura (Ítem 5). Hay que destacar la proporción amplia de estudiantes de Pedagogía que consideraron no resultarles útil el empleo de estos datos en dicho aprendizaje (41.3%) (Tabla 8). Ello puede deberse a que una buena parte de la muestra correspondiente a este grado se encontraba en primer curso y, por tanto, no han podido comprobar todavía la utilidad de los datos en el contexto de otros aprendizajes.

Tabla 9. Utilización de los datos como recurso para el aprendizaje de otra asignatura.

	G-CRI (N= 165)	G-ES (N=3)	G-ID (N=18)	G-LLCR (N=13)	G-P (N=111)	X ²	p
Sí	73.5%	-	7.4%	2.9%	16.2%	17.64	.000
No	47.5%	1.2%	5.4%	4.5%	41.3%		

De acuerdo con lo afirmado en la pregunta anterior, los alumnos de Criminología estimaron conveniente el uso de los datos como fuente de información en otras asignaturas (53.6%), junto a los de Información y Documentación (6.1%) y Lenguas,

Literatura y Cultura Románicas (4.1%) (Ítem 7) (Tabla 9). Destacar además que, si bien los alumnos de Pedagogía no señalaron especialmente significativo el uso de los datos como un recurso para aprender en otras asignaturas (Ítem 5), sí lo han valorado como fuente de información en distintas materias (35.6%). Este aspecto se puede explicar porque el empleo de los mismos en este sentido se haya evidenciado ya en algunas asignaturas recibidas durante el primer curso en esta titulación, como es el caso de Metodología de la Investigación Educativa. Y, recordemos, que era en este curso en donde se situaba una parte importante de la muestra correspondiente a este grado.

Tabla 10. Conveniencia de utilizar el uso de los datos como fuente de información en otras asignaturas.

	Criminología (N= 165)	Educ. Social (N=3)	Informac. y Documentación (N=18)	Lenguas, literatura y Culturas Románicas (N=13)	Pedagogía (N=111)	X²	p
Sí	53.6%	0.7%	6.1%	4.1%	35.6%	52.69	.000
No	46.2%	-	-	1.6%	13%		

Considerando el curso, la utilización de los datos como medio para el aprendizaje de otra asignatura (ítem 5) no es tenido en cuenta por el 24.8% del alumnado de primero, ni por el 19.8% de tercero. Son datos que difieren al considerar los restantes cursos, que se decantaron en mayor medida por su uso, tal como se evidencia en la tabla 10. En el apartado “c” de esta misma cuestión, se les preguntó por el tipo de fuentes de datos que empleaban con anterioridad al proceso formativo desarrollado. Las diferencias estriban en la utilización de las fuentes de información estadística, que fueron seleccionadas prioritariamente por el grupo de segundo curso (71.8%), el portal institucional de datos abiertos, por parte del alumnado de primero (28.6%); y, finalmente, los repositorios de datos, usados más por los alumnos de cuarto curso durante la realización de la actividad (44.4%), frente a otros recursos. La escasa formación recibida en las distintas titulaciones sobre la temática que nos ocupa, a excepción de la acción formativa realizada con motivo del Proyecto de Innovación; y la necesidad de profundización en relación con esta temática, subsanable con una mayor dedicación temporal a los distintos contenidos, son aspectos que pueden explicar los significativos porcentajes de respuestas negativas a esta cuestión, a pesar de las diferencias entre cursos.

Tabla 11. Utilización de los datos como medio para el aprendizaje de otra asignatura y fuentes de datos empleados.

	Primer curso (N=63)	Segundo curso (N=140)	Tercer curso (N=54)	Cuarto curso (N=51)	X²	p
Sí	7.4%	60.3%	8.8%	23.5%	18.53	.001
No	24.8%	40.9%	19.8%	14.5%		
Buscadores generales de datos	10.8%	40.5%	29.7%	18.9%	44.32	.001
Fuentes de información estadística	10.3%	71.8%	5.1%	12.8%		
Portal Institucional de Datos Abiertos	28.6%	57.1%	7.1%	7.1%		
Repositorio de datos	11.1%	16.7%	27.8%	44.4%		

La opinión de los estudiantes sobre si la formación y uso de datos contribuyen a desarrollar su espíritu crítico es prácticamente paritaria. El 47.2% creen que sí favorece esta opción, y consideran mayoritariamente que de esta manera pueden adquirir competencias para evaluar la información obtenida y poder seleccionar la más fiable. Estas respuestas guardan relación con los aspectos en los que les gustaría profundizar más a los encuestados, ya que al 42% les interesa la identificación de fuentes confiables de datos, y al 40%, cómo analizar y gestionar datos. No se han encontrado diferencias significativas atendiendo al espíritu crítico que puede generar la formación y el uso de datos, según la titulación, el curso y el género. Se estima, por tanto, que los estudiantes universitarios son ya más conscientes de la manipulación que pueden ejercer quienes controlan los datos y, en consecuencia, una mayor formación y empleo de datos son estimados positivamente como aspectos que van a desarrollar y/o consolidar su conciencia crítica.

Por último, en el ítem 10 se les preguntaba si consideran que la Universidad de Salamanca está trabajando para que los estudiantes tengan competencias digitales. Un 42.6% de los estudiantes aseguraron que sí se está llevando a cabo, frente al 28.4% que no lo considera, mientras que el resto no saben si se está realizando esta iniciativa. Hay que destacar el 52.3% de respuestas afirmativas por parte de los alumnos de Criminología y el 64.4% de indecisas. Por su parte, en Pedagogía (45.5%) y en Lengua, Literatura y Cultura Románicas (9.1%) se consideró con más énfasis, que no se está trabajando en este sentido por parte de la Universidad, o bien no tienen información al respecto, opción que recopila un importante porcentaje de las opiniones de toda la muestra.

Tabla 12. La Universidad de Salamanca está trabajando para que los estudiantes tengan competencias digitales.

	G-CRI (n= 165)	G-ES (n=3)	G-ID (n=18)	G-LLCR (n=13)	G-P (n=111)	X²	p
Sí	52.3%	-	9.1%	0.8%	37.9%		
No	43.2%	1.1.%	1.1%	9.1%	45.5%	27.47	.001
No sé	64.4%	2.2%	5.6%	4.4%	23.3%		

El alumnado de segundo fueron los que señalaron mayormente no tener información en relación con esta posibilidad (60%). Por tanto, en este ítem los estudiantes respondieron fundamentalmente de forma negativa, o bien indecisa, una conclusión en cierto modo normal si desde las distintas Titulaciones no se les ha informado específicamente sobre la labor que puede estar desarrollando esta Universidad.

Tabla 13. La Universidad de Salamanca está trabajando para el desarrollo de Competencias digitales.

	Primer curso (N=63)	Segundo curso (N=140)	Tercer curso (N=54)	Cuarto curso (N=51)	X²	p
Sí	28%	49.2%	12.9%	9.8%		
No	12.5%	23.9%	33%	30.7%		
No sé	18.9%	60%	8.9%	12.2	54.17	.00

4. Conclusiones

Según Coughlan (2020), la limitada investigación, existente hasta la fecha, sobre la utilidad de los datos abiertos como material para el aprendizaje sugiere una falta de conciencia o coordinación y, sobre todo, una falta de conocimiento por parte de los docentes. Sin embargo, la formación en alfabetización en datos centrada en prácticas educativas basadas en experiencias tiene repercusiones positivas para el profesorado de cualquier rama de conocimiento y su práctica profesional. Esto es así porque, a través de la formación en alfabetización en datos, el profesorado amplía sus recursos pedagógicos y mejora sus expectativas y seguridad profesional para tomar decisiones informadas e introducir cambios en sus prácticas docentes fundamentados en evidencias (Rubio et al., 2018; Campillo et al., 2019; Perinés & Campaña, 2019). Y es que cuanto más dependamos de los datos más necesario será pensar críticamente sobre lo que los datos quieren decirnos (Richardson, 2021).

Por otra parte, la alfabetización en datos permite al estudiante una experiencia de aprendizaje diferente ya que trabajar con datos reales y en bruto supone, al mismo tiempo, un desafío y un elemento motivador ya que les acerca al mundo de la práctica profesional, lo que hace que se involucren de manera activa en el proceso de aprendizaje o que repercuta en una mejora en los resultados académicos.

Puesto que, según los resultados arrojados del estudio realizado, los discentes consideran la formación en datos útil para el aprendizaje de otras asignaturas, a pesar de las dificultades encontradas, nuestra intención es continuar la línea de trabajo iniciada profundizando en la identificación y experimentación con metodologías didácticas distintas, como las prácticas educativas basadas en evidencias, que permitan trabajar la formación en datos; e incorporándola como una acción formativa transversal en los planes de estudio educativos, tal como se propone en la Carta Internacional de Datos Abiertos, adoptada en 2015, por gobiernos, sociedad civil y expertos de todo el mundo.

5. Referencias bibliográficas

- Bradbury, Alice (2018). "Datafied at four: the role of data in the 'schoolification' of early childhood education in England". *Learning, Media and Technology*, nº 1, vol. 44, pp. 7-21. <<https://doi.org/10.1080/17439884.2018.1511577>> [Consulta: 16/07/2022].
- Bradbury, Alice; Roberts-Holmes, Guy (2017). *The datafication of primary and early years education*. London, Routledge.
- Campillo, M.; Pompa, M.; Hernández, F. B.; Sánchez, M. (2019). Una experiencia de formación: la educación basada en evidencias. *Revista Digital Universitaria (RDU)*, 6 (20), 1-10. <<http://doi.org/10.22201/codeic.16076079e.2019.v20n6.a6>> [Consulta: 16/07/2022].
- Carretero, S.; Vuorikari, R.; Punie, Y. (2017). *DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use*. Union, Publications Office of the European. <<https://doi.org/10.2760/38842>> [Consulta: 24/07/2022].
- Coughlan, T. (2020). The use of open data as a material for learning. *Education Technology Research and Development*, 68, 383-411. <<https://doi.org/10.1007/s11423-019-09706-y>> [Consulta: 24/07/2022].
- CRUE Universidades Españolas (2020). *Propuesta de acciones de sensibilización para la implementación de la Agenda 2030 e inquietudes de las universidades en relación con el cumplimiento de los ODS*. <http://www.crue.org/wp-content/uploads/2021/01/Informe_Universidades_Crue-Agenda2030.pdf> [Consulta: 30/07/2022].
- Daniel, B. (2014). Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 5 (46), 904-920. <<https://doi.org/10.1111/bjet.12230>> [Consulta: 30/07/2022].
- Díaz-Arce, D.; Loyola Illescas, E. (2021). Competencias digitales en el context Covid-19: una mirada desde la educación. *Revista Innova Educación*, 1 (3). <<https://doi.org/10.35622/j.rie.2021.01.006>> [Consulta: 30/07/2022].
- European Commission (2017). *The Digital Competence framework for citizens. Luxembourg: Publication Office of the European Union*. <<https://op.europa.eu/en/publication-detail/-/publication/3c5e7879-308f-11e7-9412-01aa75ed71a1/language-en>> [Consulta: 16/07/2022].
- Gairín, J.; Ion, G. (Eds.) (2021). *Prácticas educativas basadas en evidencias. Reflexiones, estrategias y buenas prácticas*. Madrid, Narcea.
- Gómez Marcos, M^a T.; Vicente Galindo, M^a P.; Martín Rodero, H. (2019). Mujeres en la universidad española: diferencias de género en el alumnado de grado. *International Journal of Developmental and Educational Psychology*, 1 (2), 443-454.

- <<https://www.redalyc.org/jatsRepo/3498/349860126041/349860126041.pdf>>[Consulta: 16/07/2022].
- Grant, L. (2017). *Don't use professional judgement, use the actual number: The production and performance of educational data practices in an English secondary school*. [Ph.D. thesis]. University of Bristol.
- Hayes, D.; Hattam, R.; Comber, B.; Kerkham, L.; Thomson, P. (2017). *Literacy, leading and learning: Beyond pedagogies of poverty*. London, Routledge.
- Hyatt, B. (2021). A Literature Review based Conceptual Framework for Data Literacy in Construction Education. *EPiC Series in Built Environment*, 2, 487-495.
- Lewis, S. & Holloway, J. (2018). Datafying the teaching 'profession': remaking the professional teacher in the image of data. *Cambridge Journal of Education*, 1 (49), 35-51. <<https://doi.org/10.1080/0305764X.2018.1441373>> [Consulta: 24/07/2022].
- Manolev, J.; Sullivan, A.; Slee, R. (2018). The datafication of discipline: ClassDojo, Surveillance and a performative classroom culture. *Learning, Media and Technology*, 1 (44), 36-51. <<https://doi.org/10.1080/17439884.2018.1558237>> [Consulta: 16/07/2022].
- Markham, A. (2017). Troubling the concept of data in digital qualitative research, en U. Flick (Ed.), *Handbook of qualitative data collection*. London: SAGE, 511-523.
- Markham, A. (2019). Critical Pedagogy as a Response to Datafication. *Qualitative InquirI*, 8 (25), 754-760. <<https://doi.org/10.1177/1077800418809470>> [Consulta: 30/07/2022].
- O'Halloran, K. (2008). *Mathematical discourse: Language, symbolism and visual images*. New York, Continuum.
- Perinés, H.; Campaña, K. (2019). La alfabetización de los futuros docentes en investigación educativa: una reflexión teórica desde el contexto de Chile. *Revista Caribeña de Investigación Educativa (RECIE)*, 1 (3), 7-18. <<https://doi.org/10.32541/recie.2019.v3i1.pp7-18>> [Consulta: 24/07/2022].
- Pozo-Jara, A. (2017). La alfabetización informacional y la alfabetización informática: ¿iguales o diferentes?. *Dominio de las Ciencias*, 1 (3), 51-72. <<https://dominiodelasciencias.com/ojs/index.php/es/article/view/625>>[Consulta:30/07/2022].
- Raffaghelli, J. E. (2019). Datafificación y Educación Superior: Hacia la construcción de un marco para la alfabetización en datos del profesorado universitario. *Revista Interamericana de Investigación, Educación y Pedagogía*, 1 (13), 177-205. <<https://doi.org/10.15332/25005421/5466>> [Consulta: 24/07/2022].
- Richardson, S. (2021). The new physicality of data. *Business Information Review*, 2 (38), 67-74. <<https://doi.org/10.1177/02663821211020194>> [Consulta: 16/07/2022].
- Rubio, M^a J.; Torrado, M.; Quirós, C.; Valls, R. (2018). Autopercepción de las competencias investigativas en estudiantes de último curso de Pedagogía de la Universidad de Barcelona para desarrollar su Trabajo de Fin de Grado. *Revista Complutense de Educación*, 2 (29), 335-354. <<https://doi.org/10.5209/RCED.52443>> [Consulta: 30/07/2022].
- Selwyn, N.; Henderson, M.; Chao, S. H. (2015). Exploring the role of digital data in contemporary schools and schooling- "200,000 lines in an Excel spreadsheet. *British Educational Research Journal*, 5 (41), 767-781. <<https://doi.org/10.1002/berj.3186>> [Consulta: 30/07/2022].