

# Estudio de la accesibilidad de los documentos científicos en soporte digital

Marta Díaz

4all-L@b CETpD  
Universitat Politècnica  
de Catalunya  
08800 Vilanova i la  
Geltrú  
marta.diaz@upc.edu

Maite Barrios

Departament de  
Metodologia de les  
Ciències del  
Comportament.  
Universitat de  
Barcelona.  
Barcelona  
mbarrios@ub.edu

MiquelTérmens  
Mireia Ribera

Departament de  
Biblioteconomia i  
Documentació.  
Universitat de  
Barcelona. Barcelona  
termens@ub.edu,  
ribera@ub.edu

Pere Ponsa

Daniel Guasch

4all-L@b CETpD  
Càtedra Accessibilitat  
Universitat  
Politécnica de  
Catalunya  
08800 Vilanova i la Geltrú  
pedro.ponsa@upc.edu  
daniel.guasch@upc.edu

## Resumen

El artículo analiza, los problemas de accesibilidad de los artículos científicos en soporte digital. Se analizan los dos formatos más utilizados para la publicación de artículos científicos en soporte digital: HTML y PDF, estudiando la variabilidad interna que representa la presencia de sumarios o de tablas internas o vinculadas. El estudio se realizó en dos colectivos: 30 participantes ciegos y 30 participantes no ciegos. Los resultados muestran que los participantes ciegos presentan una mejor ejecución en documentos HTML que presentan ayudas en la navegación (presencia de sumario) y tablas internas. La ejecución de los participantes no ciegos parece ser independiente del tipo de formato (HTML vs PDF), aunque se beneficiarían como en el caso de participantes ciegos en la inclusión de ayudas a la navegación en el formato HTML.

## 1. Introducción

Un documento digital será usable cuando cumpla los objetivos para los que se ha diseñado de forma eficiente, eficaz y satisfactoria [4] en un contexto de uso dado, y será accesible cuando, además, estos objetivos se cumplan para personas con un amplio abanico de capacidades [5]. Dillon realizó un análisis de la usabilidad de los artículos científicos, que establece que este tipo de documento se lee básicamente por motivos laborales (para estar al día, como fuente de referencia o de aprendizaje) y que el tipo de lectura más habitual es la lectura en escaneo, la

lectura no secuencial de las principales secciones y la lectura completa del texto [1]. En cuanto a las discapacidades, las que afectan la lectura son principalmente los problemas de visión, la dislexia, los problemas de movilidad y las discapacidades cognitivas [2]. En concreto, los requerimientos de los lectores ciegos, recogidos por Petrie tras realizar diversas pruebas con usuarios y entrevistas [7] son poder disponer de una versión hablada de todo el texto, y poder controlar la navegación (rebobinar, avanzar a saltos, pausar), el volumen y la velocidad. Para cumplir estos requisitos, a menudo se requiere información adicional en el propio documento procesable por la tecnología de asistencia, como por ejemplo, la inclusión de descripciones alternativas para las imágenes [6]. La accesibilidad de los documentos digitales tiene dos vertientes. Por una parte, su contenido responde a una serie de decisiones de sus autores, por lo que respecta a arquitectura de la información, inclusión de elementos complejos o multimedia, registro lingüístico, etc. que afectaran directamente a su accesibilidad. Los documentos digitales se crean en un formato de archivo concreto, HTML y PDF, y que por su definición técnica tienen un potencial de accesibilidad que puede ser explotado en mayor o menor grado por los usuarios o por las herramientas de autor que los crean.

## 2. Objetivos del estudio

El presente estudio pretende identificar los aspectos del formato y de la arquitectura de la información de documentos científicos que

mejoran la eficiencia, eficacia y satisfacción en lectores ciegos, estudiando (según los principios de usabilidad universal) si estos aspectos inciden sobre la eficiencia, eficacia y satisfacción de usuarios sin discapacidad. Para ello se plantean las siguientes hipótesis a contrastar:

**Hipótesis 1:** El formato del documento (PDF/HTML) afecta la actividad lectora.

**Hipótesis 2:** Las ayudas a la navegación (sumario/marcadores) facilitan la actividad lectora. Se parte de la suposición que los textos con marcadores o sumarios facilitan la actividad lectora que se traducirá, como en el caso anterior, en unos resultados mejores tanto en las variables de eficiencia como eficacia y satisfacción.

**Hipótesis 3:** La división en bloques de un documento afecta la actividad lectora.

### 3. Método

La muestra estuvo compuesta por 30 participantes con ceguera total, usuarios de lectores de pantalla, (contactados gracias a la mediación de la Fundación ONCE) y 30 participantes sin discapacidad, docentes de la facultad de Biblioteconomía y Documentación de la Universidad de Barcelona.

#### 3.1. Instrumentos

##### *Entrevista*

Se elaboró una entrevista ad hoc para recoger los datos sociodemográficos del participante, así como aspectos relacionados con el uso de ordenadores y de acceso a Internet.

##### *Documentos para la tarea de lectura*

Los documentos usados en el estudio fueron elaborados atendiendo los criterios de grado de dificultad, similitud a un artículo científico, longitud y accesibilidad. Se generaron un total de 5 documentos con las siguientes características: un documento en PDF con marcadores y otro sin marcadores, un documento HTML con tablas incluidas sin sumario, un documento HTML con tablas vinculadas sin sumario y un documento HTML con tablas incluidas y con sumario.

##### *Preguntas*

Para cada uno de los documentos se generaron un total de 10 preguntas referidas a aspectos

concretos citados en los textos. Las preguntas presentaban diferente grado de dificultad e implicaban la búsqueda en diferentes partes del documento (tablas, bibliografía y texto). Se realizó una prueba piloto con 6 participantes voluntarios, para valorar la comprensión y dificultad de cada una de las preguntas. Después de diversas pruebas piloto, las 10 preguntas se redujeron a 5 (tres referidas al texto, una referida al contenido de las tablas y una referida a la bibliografía), entre las que presentaban un nivel de dificultad menor (menor tiempo de respuesta, mayor porcentaje de aciertos y menor cantidad de repeticiones de la pregunta). Se estableció un tiempo límite de 6 minutos para resolver cada una de las preguntas.

##### *Cuestionario de valoración de la tarea*

Se diseñó un cuestionario ad hoc de 12 ítems que se administró después de las tareas de búsqueda de cada uno de los documentos, con el fin de evaluar la dificultad y valoración de la tarea.

#### 3.2. Procedimiento

A cada participante se le administraron los 5 documentos y las respectivas preguntas. Se controló el efecto de orden y de aprendizaje de la tarea balanceando tanto el orden de presentación de los documentos como el de las preguntas. Para cada pregunta realizada se tomaron medidas de tiempo y precisión (acierto/error). Los participantes ciegos contaron con el programa Jaws for Windows V. 6.20, de Freedom Scientific. El tiempo de evaluación osciló entre una hora y dos horas y media. Las variables sociodemográficas se presentan mediante estadísticos descriptivos. Para los contrastes de hipótesis se utilizó la prueba de t-student para muestras relacionadas o bien la prueba de McNemar. En el caso de no cumplimiento de las condiciones de aplicación de la prueba de McNemar se procedió a realizar una prueba de bondad de ajuste. Los contrastes fueron considerados estadísticamente significativos cuando  $p < 0,05$ .

### 4. Resultados

A continuación se presenta la valoración de resultados agrupados por participantes.

#### 4.1. Participantes no ciegos

La edad promedio se situó en 42,1 años (Mediana (Md)= 43 años, desviación típica (DT)= 9,06 años). Todos tenían estudios universitarios finalizados y usaban el ordenador diariamente. El promedio de uso del ordenador fue de 17,4 años (Md= 18 años y DT= 5,06).

##### *Formato del documento (PDF vs HTML)*

En cuanto al tiempo de resolución de las tareas de búsqueda, no se observaron diferencias estadísticamente significativas al comparar los documentos en PDF y HTML. La precisión en la respuesta fue similar en ambos documentos, observándose siempre más de un 90% de aciertos para cada pregunta. Los datos no mostraron diferencias para ninguno de los 12 ítems del cuestionario de valoración de la tarea.

##### *Ayudas a la navegación. Formato PDF*

No se observaron diferencias estadísticamente significativas para los documentos PDF con y sin ayudas a la navegación (marcadores) para ninguna de las variables de tiempo de resolución, precisión y valoración de la tarea.

##### *Ayudas a la navegación. Formato HTML*

El tiempo de resolución en los documentos HTML con y sin ayudas a la navegación (sumarios) mostró diferencias estadísticamente significativas en dos de las cinco preguntas (pregunta 1 referida a texto y pregunta 5 referida a la bibliografía). La resolución fue más rápida cuando el documento presentaba sumario. No se hallaron diferencias en la precisión de la respuesta, observándose un alto porcentaje de respuestas correctas (superior al 90%). Ninguno de los ítems del cuestionario mostró diferencias estadísticamente significativas.

##### *División en bloques. Formato HTML: Tablas incluidas y vinculadas*

No se observaron diferencias estadísticamente significativas para ninguna de las variables de tiempo de resolución, precisión y valoración de la tarea entre los dos tipos de documentos.

#### 4.2. Participantes ciegos

La media de edad fue de 36,57 años (Md= 36, DT= 10,6 años). El 43,3% tenía nivel de estudios

Tabla 1. HTML con y sin sumario en la variable tiempo (en segundos).

		Media (DT)	t (gl)	P	R
P. 1	HTML-S	28,90 (28,92)	-2,23 (29)	0,03	0,38
	HTML-NS	54,13 (61,78)			
P. 5	HTML-S	20,20 (19,84)	-2,35 (29)	0,03	0,40
	HTML-NS	42,63 (51,12)			

HTML-S : HTML con sumario, HTML-S: HTML sin sumario, DT: Desviación típica, t: t de student de medidas repetidas. GL: grados de libertad. p: significación estadística, r: tamaño del efecto.

Tabla 2. Comparaciones entre PDF y HTML en la variable tiempo (en segundos)

		Media (DT)	t (gl)	P	R
P. 2	PDF	161,17 (114,97)	2,38 (29)	0,02	0,40
	HTML	100,10 (86,71)			
P. 4	PDF	138,30 (110,02)	4,04 (29)	<0,001	0,6
	HTML	56,33 (47,01)			
P. 5	PDF	137,30 (125,65)	2,30 (29)	0,03	0,39
	HTML	82,60 (80,12)			

DT: Desviación típica, t: t de student de medidas repetidas. gl: grados de libertad., p: significación estadística, r: tamaño del efecto.

universitarios, un 40% estudios secundarios y un 16,7% había finalizado estudios primarios. Un 96,6% afirmaron usar el ordenador y de estos un 90% lo usaba a diario. El promedio de uso del ordenador fue de 10,23 años, (Md = 10 años y DT= 5,39 años) y el uso de lectores de pantalla de 6,2 años (Md= 6 años y DT= 1,77 años).

##### *Formato del documento (PDF vs HTML)*

En el tiempo de resolución de las tareas de búsqueda se observaron diferencias estadísticamente significativas (Tabla 2) en tres de las preguntas referidas a diferentes aspectos del documento: pregunta 2 (texto), 4 (tabla) y 5 (bibliografía). En todos los casos, el tiempo de búsqueda fue menor en los documentos HTML.

Los aciertos fueron estadísticamente diferentes, en la pregunta 2 (texto) ( $\chi^2 = 80,23$ ,  $p < 0,001$ ), pregunta 3 (texto) (McNemar-Bowker = 9, g.l. = 3,  $p = 0,029$ ) y pregunta 4 (tabla) ( $\chi^2 = 26,69$ ,  $p < 0,001$ ). En todos los casos, la precisión fue superior en los documentos con formato en HTML. Respecto al cuestionario, se observan diferencias estadísticamente significativas en 8 de los 12 ítems. En todos los casos, los documentos en formato HTML, mostraron una percepción más favorable.

#### *Ayudas a la navegación. Formato PDF*

El tiempo de resolución en los documentos PDF con y sin ayudas a la navegación (marcadores) fue significativamente más rápida en la pregunta 1 (texto) (t-student = 2,63,  $p = 0,01$ ,  $r = 0,45$ ), cuando no se disponía de ayudas en la navegación (PDF sin marcadores). No se observaron diferencias en cuanto a la precisión de la respuesta ni en ninguno de los ítems del cuestionario.

#### *Ayudas a la navegación. Formato HTML*

Se hallaron diferencias estadísticamente significativas en el tiempo de ejecución tan sólo en la pregunta 4 (tabla) (t-student: 3,16,  $p = 0,003$ ,  $r = 0,51$ ). La ejecución en esta pregunta, fue más rápida cuando se disponía de sumario. Los datos no mostraron diferencias en cuanto a la precisión de la respuesta ni en ninguno de los ítems del cuestionario.

#### *División en bloques. Formato HTML: Tablas incluidas y vinculadas*

Cuando se realizó una pregunta sobre las tablas el tiempo en ejecución fue significativamente menor cuando el documento presentaba las tablas incluidas (Preg. 4: HTML con tablas incluidas: media = 56,33, DT = 47,01, HTML con tablas vinculadas: media = 103,30, DT = 88,65,  $t = -2,7$ , g.l. = 29,  $p = 0,01$ ,  $r = -0,45$ ). No se observaron otras diferencias estadísticamente significativas para el resto de preguntas en la variable tiempo, precisión, ni para ninguno de los ítems del cuestionario.

## **5. Conclusión**

Según los objetivos planteados se han estudiado las características de accesibilidad que presentan los dos formatos técnicos más utilizados en revistas electrónicas (HTML y PDF), así como

una parte de la variabilidad con la que estos formatos pueden ser presentados.

Los resultados obtenidos en este estudio muestran que los participantes ciegos presentan una mejor ejecución en documentos con formato HTML que presentan ayudas en la navegación (presencia de sumario) y las tablas se incluyen en el documento. Las ayudas en la navegación en el caso de documentos PDF, parece dificultar la tarea más que facilitar, aunque estos resultados podrían deberse a una menor familiaridad con el formato y sus aplicaciones. La ejecución de los participantes no ciegos parece ser independiente del tipo de formato (HTML vs PDF), aunque se beneficiarían como en el caso de participantes ciegos en la inclusión de ayudas a la navegación en el formato HTML.

## **Agradecimientos**

Este trabajo ha sido subvencionado por el Programa de Estudios y Análisis, referencia EA2006-0049 del Ministerio de Educación, Cultura y Deporte.

## **Referencias**

- [1] Dillon, A. (2004). Designing usable electronic text (2nd ed.). Boca Raton: CRC Press
- [2] Dix, A. Human-computer interaction. Pearson Education, 2004. 834 p.
- [3] Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated. Conference on Human Factors in Computer Systems, The Hague
- [4] ISO 9241-11. (1998) Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability
- [5] ISO/TS 16071 (2003) Ergonomics of Human-System Interaction. Guidance on Accessibility for Human-Computer Interfaces
- [6] Paciello, M. G. (2000). Web accessibility for people with disabilities. Lawrence, Kansas: CMP Books
- [7] Petrie, H. [et al.] (2002). Universal interfaces to multimedia documents. IEEE International Conference on Multimodal Interfaces (4th). 319-324