

Resumen de la Reunión del GT Google, celebrada el 16 de abril de 2007

Orden del Día:

1. Auditoría de colecciones de la BUC: informe de Google.
2. Guía de digitalización UCM-Google: finalización
3. Sigüientes pasos del Proyecto.
4. Ruegos y preguntas

1. Informe sobre la auditoría de colecciones de dominio público realizado por Google.

Se comenta el informe remitido por Google que se encuentra en la intranet a disposición del grupo de trabajo. Los resultados del informe se consideran muy positivos para la biblioteca. Habrá una 100.000 obras susceptibles de digitalización. Los principales motivos de exclusión de obras se deben al deficiente estado de encuadernación y a que se sobrepasan las dimensiones físicas capaces de ser gestionadas por las estaciones de digitalización de Google.

Se considera que el informe es un excelente punto de partida para realizar un Plan de Encuadernación, tanto a nivel de la Biblioteca General como en la F. de Psicología, que permita incrementar el número de obras a digitalizar. A este respecto, se estima que podría ser viable la solicitud de una partida financiera "ad hoc" de 100000 € vía Consejo Social o Fundación UCM. Con todo, Javier Tacón informa que aquellas obras cuya exclusión ha sido motivada por el escaso ángulo de apertura que permite la encuadernación seguirán sin poder digitalizarse, por lo que habrá que efectuar este trabajo con otro proveedor.

La Biblioteca Histórica presenta una propuesta de envasado al vacío de libros del fondo antiguo que tiene por finalidad mantener las condiciones medioambientales de las obras. Sería interesante que Google también lo utilizara.

En este punto, José Antonio Magán informa que la Gerencia de la UCM ha autorizado la creación de una bolsa de trabajo para contratar a dos restauradores de personal laboral.

2. Guía de digitalización UCM-Google: finalización

La BH se compromete a tener finalizado los últimos flecos pendientes de la Guía para presentarla a Google en la próxima reunión.

3. Plan de Logística. Las acciones a efectuar son:

Se resumen los principales temas pendientes de resolver:

- Fijar el orden por centros en que va a realizarse la digitalización.
- Determinar los problemas de accesibilidad en los centros.
- Dejar resuelto al tema de los metadatos: los metadatos generales de la BUC y los metadatos correspondientes al Manifest (la lista de libros que van en cada cada).

Tareas:

- Preguntar por qué en Psicología se repite la cifra 4,35%:
 - Respuesta de Google: *We checked the data, and it looks like our reports are accurate. There were only 2 books at the Psicologia that had problems making them unsuitable for scanning, and both of the books were in very bad condition, and had problems with all four of the characteristics that we were evaluating. That is why they show the same percentages across the board. The Veterinaria and Geografía libraries show 0% across the board, because all of the books we surveyed there were suitable for scanning.*
- Sacar los metadatos de Dioscóridos: *Está previsto enviar los metadatos (ya se enviaron las imágenes) a finales de esta semana.*
- Solicitar a Google que nos remitan su protocolo de condiciones medioambientales en el Scanning Center:

- **Respuesta de Google:** 1) The Google protocol of the environmental conditions of the scanning center (humidity data, temperature, etc.). We do not monitor the humidity of our scanning center. The air conditioning is set for a temperature of 72 degrees Fahrenheit (22.22 degrees Centigrade). This keeps operators comfortable and keeps the air conditioning running to offset heat generated by the computers and station lights. Typically, the books are in this environment for a relatively short time (a few days).
- Solicitar a Google las características técnicas del camión de transporte empleado para digitalización:
 - **Respuesta de Google:** The technical features of the trucks for the books daily shipments (how is the inner part, etc.).
Transport – The transit route from the scanning center to and between the libraries and back to the scanning center will be planned in advance and normally follow this pre-planned route. Alternate routes will also be pre-planned and in the event of conditions that prevent or delay the shipment, the driver will have the authority to utilize such alternate routes.

At campus libraries - There will be a pre-defined pick-up time which correlates with the Library and Google schedule. At each respective library the carts for that library are first unloaded and staged in the designated library shipping/receiving area and the new carts picked up and loaded in the truck. The driver will assure compliance with all parties involved on pick-up times and be flexible to accommodate last minute changes in schedules. At no time will the driver leave the truck unattended without securing the truck. We use standard local transport trucks. The delivery truck utilized for daily runs typically will have specifications similar to these:
 - Size – 16 to 24 Feet
 - Height of Floor - Dock Height
 - Transmission – Auto or Manual
 - Lift gate - Yes
 - Air-Ride Equipped - Yes
 - Fuel type – Gasoline or Diesel
 - Cargo Tie Down – Etrack Tie-Down System
- Solicitar a Google las especificaciones del atril que utilizan.
 - **Respuesta de Google:** As mentioned, the cradles are made of painted wood, and are open to an angle of 120 degrees. They can be moved independently to accommodate the width of a book's spines, and are designed so that the spine hangs freely, and the book rests on its boards on the cradle.
- Consultar si van a dar cabida a obras de gran tamaño.
 - **Respuesta de Google.** Los libros de gran formato van a poder digitalizarse en el futuro porque van a disponer de escáneres adecuados a libros de mayor tamaño. Por ello, cuando hagamos la selección deberemos anotar las medidas de esos libros que en este momento no son aptos para su digitalización.
- Comunicar a Google la pregunta del Rector respecto a si se van a digitalizar las obras selectas en la BH.
 - **Respuesta de Google.** Currently, we don't have a system designed to scan books in very small spaces, such as the Biblioteca Historica. All of our

development to this point has been with the goal of large-scale digitization, which requires significant power and computing infrastructure.

I think that we would very much like to digitize the materials in the Histórica facilities, but currently our focus is on the more easily accessible materials. Given that there are still many many of these types of materials to do, I imagine that our focus will remain on high-volume digitization for some time to come (years).

So I don't want to give you a false impression of our priorities and our thinking, or commit to a time frame that we are not prepared to meet. We would like to digitize these materials, and think that it is worthwhile to do so, but realistically would not be in a position to start thinking about how to do these for a year or two.