Cómo abordar el estudio de los manuscritos griegos desde las nuevas tecnologías

Carmen García Bueno

23/4/25 carmen.garcia.bueno@cchs.csic.es



↑ CAAL (Centro per l'Automazione dell'Analisi Letteraria), Gallarate (IT), 1967) Trabajadoras con tarjetas perforadas para el *Index Thomisticus*.



↑→ Personal investigador del proyecto Diccionario Griego-Español (Instituto de Filología (actual ILC), CSIC). Antes de 2008.



divulgación licencia Creative Commons open source redes sociales publicación en

abierto herramientas digitales financiación interactividad IA copias de seguridad evaluación académica docencia preservación nuevas formas de investigar ciencia en abierto bibliografía memoria justificativa

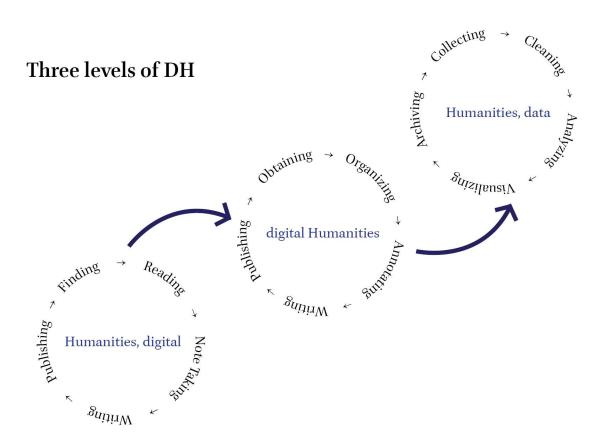
¿Humanidades Digitales?

Nivel 1: las Humanidades, pero hechas de manera digital.

Nivel 2: *Humanidades Digitales* propiamente dichas.

Nivel 3: Humanidades de naturaleza computacional.

Este tercer nivel puede sonar atractivo, pero no es algo que puedas alcanzar fácilmente la mayoría de las veces. Para campos como los nuestros, los dos primeros niveles son mejores.



van Lit, C. Three levels of DH. 01/10/2021. https://digitalorientalist.com/2021/10/01/how-to-level-up-in-dh/

Humanidades Digitales: el ordenador es tu amigo

Herramienta con una capacidad para analizar conjuntos de datos de gran volumen según unos parámetros dados y producir con ello un resultado de manera mucho más rápida de lo que puede un humano.

Ese resultado (salida, *output*) **no es necesariamente** ni la solución definitiva ni la mejor. Nosotros seguimos teniendo la última palabra.

El investigador (de HASS¹) actual necesita:

- Entender (mínimamente) cómo funciona el ordenador
- Usar herramientas más especializadas
- Saber qué están haciendo otros

- → formación
- → repositorios de herramientas y servicios
- *→* ¿?

¹ HASS = Humanities, Arts and Social Sciences, por oposición a STEM (Science, Technology, Engineering and Mathematics)

Humanidades Digitales: el ordenador es tu amigo

Formación:

- Programming Historian (https://programminghistorian.org/)
- #dariahTEACH (https://teach.dariah.eu/)
- algún recurso específico (p. ej. TTHub: https://tthub.io/)
- en paleografía griega (*Greek Paleography Thematic Pathways DVL*: https://spotlight.vatlib.it/greek-paleography)

Herramientas:

- SSH Open Market > Service catalogue (https://www.sshopencloud.eu/service-catalogue)
- Dariah-EU Tools and Services Catalogue (https://www.dariah.eu/tools-services/tools-and-services/)
 - OpenMethods (https://openmethods.dariah.eu/)
- Awesome Digital Humanities (https://dh-tech.github.io/awesome-digital-humanities/)
- Digital-Humanities-Toolkit (https://github.com/pacian/Digital-Humanities-Toolkit)
- X TAPOR^3 X DH Toy Chest X DIRT ? AlternativeTo

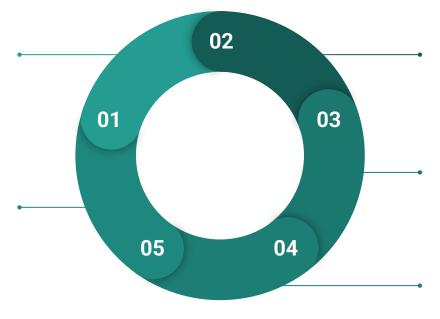
¿Qué hay de nuevo, viejo?

Digitalización: del objeto físico al electrónico

Repositorios. Proyectos de digitalización. Técnicas no intrusivas para conseguir ver más.

Divulgación, preservación et al. (et IA?)

¿Dónde pongo esto? Mi trabajo, para todos los públicos. Otros recursos útiles. Nulla aesthetica sine ethica: el derecho a cuestionarse la IA.



Transcripción: ¡a leer!

OCR, HTR, ATR. Etiquetado. El texto "objeto" (paleografía). El texto "texto" (filología).

Edición digital (o no)

Colación. Finalidad o complemento. (¿) Límites de la ED. (?)

"Lo que no es texto"

Iluminación. Encuadernación. Filigranas y otros aspectos del soporte. Bibliotecas, lectores y posesores.

1. Digitalización: del objeto físico al electrónico

Repositorios

→ - <u>Pinakes</u>

- <u>NTVMR</u>
- <u>CSNTM</u>
- <u>Álbum de copistas de manuscritos griegos en España</u>

 Proyectos de digitalización (metadatos)

- <u>DigiVatLib</u> (2010-), con la <u>Polonsky Foundation</u> (2012-2018)
- <u>Cambridge and Heidelberg</u> (Polonsky Foundation, 2019-)
- <u>HMML</u>
- <u>DigitES C</u> (ILC, CSIC) + <u>RBMEDigital</u>
- Técnicas no intrusivas para conseguir ver más →
- Proyectos sobre palimpsestos:
 - <u>Sinai Palimpsests Project</u> (EEUU)
 - <u>Greek Palimpsests @ ÖNB</u> (Austria)
 - <u>Palimpsests and Imaging</u> (Francia)
 - Para saber más:
 - Vatican Palimpsests Thematic Pathways DVL
 - "Projects" Early Manuscripts Electronic Library

2. Transcripción: ¡a leer!

- OCR (Optical Character Recognition)
 HTR (Handwritten Text Recognition)
 ATR (Automatic Text Recognition)
 Etiquetado (XML(-TEI))
- El texto "objeto" (paleografía)

- Software generalista (Google Drive, ChatGPT y similares, ABBY OCR, OCR4all, EasyOCR, PaddleOCR, Gimp + AirTable...)
 - Software especializado en transcripción (<u>Transkribus</u>, <u>eScriptorium</u>, <u>Transkriptorium</u>, <u>Kraken</u>, <u>Monk</u>, <u>Aletheia</u>, <u>Tesseract</u>, <u>Calfa</u>, <u>ArkIndex</u>...)
 - Software especializado en anotación de imágenes (<u>DigiPal</u>, <u>IIIF</u> + <u>Mirador</u>, <u>OpenAnnotation</u> + <u>Annotorius</u>...)
- \rightarrow <u>CDDGB</u> (ss. I a.C.-IX)
 - Análisis de colofones:
 - Σημειώματα-Κώδικες (base de datos de colofones de mss griegos ss. IX-XV)
 - Proyecto EPPE (ILC, CSIC)
 - Ommundsen, Å., Conti, A.K., Haaland, Ø.A. *et al.* How many medieval and early modern manuscripts were copied by female scribes? A bibliometric analysis based on colophons. *Humanit Soc Sci Commun* **12**, 346 (2025). https://doi.org/10.1057/s41599-025-04666-6
- PLN, POS, estilometría, análisis de sentimientos, redes de contactos, lugares citados...

El texto "texto" (filología)

lo que decía el original \rightarrow

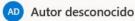
στολήν Πορφυρίου cod. (ed. G. Parthey, Berolini 1857, p. 1-294; ed. recenter M. Sicherl, Berlin 1956); praemittitur scholium Ἰστέον ὅτι ὁ φιλόσοφος Πρόκλος—καὶ ἱστορίσαντα (ed. P. Henry, Études Plotiniennes I: Les États du texte de Plotin, Bruxellis 1938, p. 284, N°. R; cf. Plotini opera, ed. Fr. Creuzer, I, Oxonii 1835, p. XXXI adn. 2).

↓lo que leyó ABBY

Copiado por tres diversas manos: (ff. VI-VII. 1-44*) por Manuel Glynzunios; (ff. 45-170*) letra gruesa y vertical; (ff. 173-247*) letra fina e inclinada hacia la derecha. Consta de 32 cua-

(ed. N. Festa, Leipzig 1891, p. 3-99); 4 (ff. 128v-170v) in Niccomachi arithmeticam introductio, ἤτις περιέχει τῆς ἀπλῶς μαθηματικῆς διδασκαλικῶς τοὺς ἀρχικοὺς λόγους add. cod. (ed. H. Pistelli, Leipzig 1894, p. 3-125). 2 (ff. 173-247v) liber de mysteriis>, εἰς τὴν ἐπιστολὴν Πορφυρίου cod. (ed. G. Parthey, Berolini 1857, p. 1-294; ed. recenter M. Sicherl, Berlín 1956); praemittitur scholium Ἰστέον ὅτι ὁ φιλόσοφος Πρόκλος — καὶ ἰστορίσαντα (ed. P. Henry, Études Plotiniennes I: Les États du texte de Piolín, Bruxellis 1938, p. 284, N°. R; cf. Plotini opera, ed. Fr. Greuzer, I, Oxonii 1835, p. XXXI adn. 2).

Copiado por tres diversas manos: (ff. VI-VII. 1-44v) por Manuel Glynzunios; (ff. 45-170v) letra gruesa y vertical; (ff. 173-247v) letra fina e inclinada hacia la derecha. Consta de 32 cuaterniones



sí, debería ser Plotin, pero lo dejo porque me ha alegrado la mañana

2. Transcripción: ¡a leer!

- OCR (Optical Character Recognition)
 HTR (Handwritten Text Recognition)
 ATR (Automatic Text Recognition)
 Etiquetado (XML(-TEI))
- El texto "objeto" (paleografía)

- Software generalista (Google Drive, ChatGPT y similares, ABBY OCR, OCR4all, EasyOCR, PaddleOCR, Gimp + AirTable...)
 - Software especializado en transcripción (<u>Transkribus</u>, <u>eScriptorium</u>, <u>Transkriptorium</u>, <u>Kraken</u>, <u>Monk</u>, <u>Aletheia</u>, <u>Tesseract</u>, <u>Calfa</u>, <u>ArkIndex</u>...)
 - Software especializado en anotación de imágenes (<u>DigiPal</u>, <u>IIIF</u> + <u>Mirador</u>, <u>OpenAnnotation</u> + <u>Annotorius</u>...)
- \rightarrow <u>CDDGB</u> (ss. I a.C.-IX)
 - Análisis de colofones:
 - Σημειώματα-Κώδικες (base de datos de colofones de mss griegos ss. IX-XV)
 - Proyecto EPPE (ILC, CSIC)
 - Ommundsen, Å., Conti, A.K., Haaland, Ø.A. *et al.* How many medieval and early modern manuscripts were copied by female scribes? A bibliometric analysis based on colophons. *Humanit Soc Sci Commun* **12**, 346 (2025). https://doi.org/10.1057/s41599-025-04666-6
- PLN, POS, estilometría, análisis de sentimientos, redes de contactos, lugares citados...

El texto "texto" (filología)

3. Edición digital (o no)

- Colación
- Finalidad o complemento
- (¿) Límites de la ED (?)

- → Software generalista: hojas de cálculo, texto plano con tabuladores, MarkDown...
 - Software especializado: <u>CTE</u>, <u>Juxta</u>, <u>ChrysoCollate</u>, (XML-)TEI + (XSLT+XPath+XQuery+XHTML) (<u>oXygen</u>, <u>VSC</u>+extensions)...
- \rightarrow ¿Qué es una ED(A)¹?
 - Mancinelli, T., y E. Pierazzo. *Che cos'è un'edizione scientifica digitale*. Roma: Carocci editore, 2020.
 - Sahle, P. «What Is a Scholarly Digital Edition?» En *Digital Scholarly Editing*, editado por M. J. Driscoll y E. Pierazzo, 19-39. *Theories and Practices* 4. Cambridge: Open Book Publishers, 2016. https://doi.org/10.11647/0BP.0095.
 - Catálogos de ED(A)s:
 - Franzini, G. «Catalogue of Digital Editions». Catálogo en línea, 2015-. https://dig-ed-cat.acdh.oeaw.ac.at/.
 - Sahle, P. «A Catalog of Digital Scholarly Editions». Catálogo en línea, 2020-. https://www.digitale-edition.de/exist/apps/editions-browser/index.html.
 - <u>EVT</u> (2, <u>3</u>), <u>TEI Publisher</u> (+ eXistDB), PHP (diseño *ad hoc*), LaTeX...
 - DBBE, ParaTexBib

 $^{^{1}}$ ED(A) = Edición Digital (Académica)

4. "Lo que no es texto"

lluminación estudio/reconstrucción del paisaje sonoro y de la arquitectura (texto+imagen+arqueología): Proyecto MABILUS: Rodríguez Suárez, A. "Semantra and Bells in the Skylitzes Matritensis". MABILUS, I Simposio Internacional. Madrid, 13 de junio de 2024 (ver) Calahorra Bartolomé, A. Keleusate : arquitectura, arte y ceremonia en la gran palacio de Constantinopla. Tesis doctoral. Universidad Complutense de Madrid, 26/5/2023. https://hdl.handle.net/20.500.14352/92817 (más) Saint George on a Bike Encuadernación proyecto <u>Ligatus</u> (UAL, UK) proyecto <u>Aracne</u> (Museo del Prado) representación de la estructura: $\underline{\text{VisColl}}$ ($\rightarrow \underline{\text{EVT}}$), $\underline{\text{StruViMan}}$ ($\rightarrow \text{Pinakes}$) Filigranas y otros aspectos del soporte Berstein: The memory of paper Briquet online Reconstrucción virtual de la biblioteca de Matías Corvino Bibliotecas, lectores y posesores Reconstrucción de la biblioteca de Michel de Montaigne





← Skylitzes Matritensis, f. 28r (detalle): monjes en un monasterio, estudiando, mientras otro llama con un semantra

Skylitzes Matritensis, f. 55r (detalle) ilustración de un evento en el hipódromo, en el que se puede ver al emperador en el *kathisma*. A la derecha, reconstrucción de Calahorra Bartolomé 2023



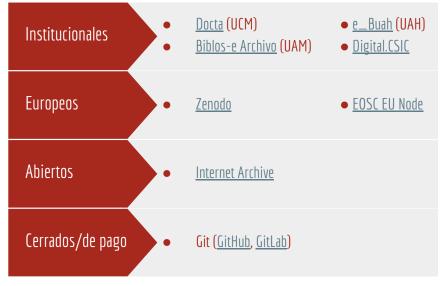
Lám. 3 El Kathisma tras la reforma de Justiniano.

4. "Lo que no es texto"

lluminación estudio/reconstrucción del paisaje sonoro y de la arquitectura (texto+imagen+arqueología): Proyecto MABILUS: Rodríguez Suárez, A. "Semantra and Bells in the Skylitzes Matritensis". MABILUS, I Simposio Internacional. Madrid, 13 de junio de 2024 (ver) Calahorra Bartolomé, A. Keleusate : arquitectura, arte y ceremonia en la gran palacio de Constantinopla. Tesis doctoral. Universidad Complutense de Madrid, 26/5/2023. https://hdl.handle.net/20.500.14352/92817 (más) Saint George on a Bike Encuadernación proyecto <u>Ligatus</u> (UAL, UK) proyecto <u>Aracne</u> (Museo del Prado) representación de la estructura: $\underline{\text{VisColl}}$ ($\rightarrow \underline{\text{EVT}}$), $\underline{\text{StruViMan}}$ ($\rightarrow \text{Pinakes}$) Filigranas y otros aspectos del soporte Berstein: The memory of paper Briquet online Reconstrucción virtual de la biblioteca de Matías Corvino Bibliotecas, lectores y posesores Reconstrucción de la biblioteca de Michel de Montaigne

5. Divulgación, preservación *e.a.* (*et IA*?)

- ¿Dónde pongo esto? \longrightarrow - Repositorios



Alojamiento en servidores de la institución o externos (gasto a proyecto o personal)

5. Divulgación, preservación *et al.* (*et IA*?)

Otros recursos útiles

→ - Bibliografía (<u>ejemplo</u>)

- Colecciones digitales: Omeka (ejemplo)

Mapas: Recogito (ejemplo), líneas de tiempo (<u>Timelinejs</u>)

- Mi trabajo, para todos los públicos

- Principios FAIR + Ciencia Abierta (MICIIN, UNESCO)

- Findable (Encontrable)

- Accessible (Accesible)

Interoperable (Interoperable)

- Reusable (Reutilizable)

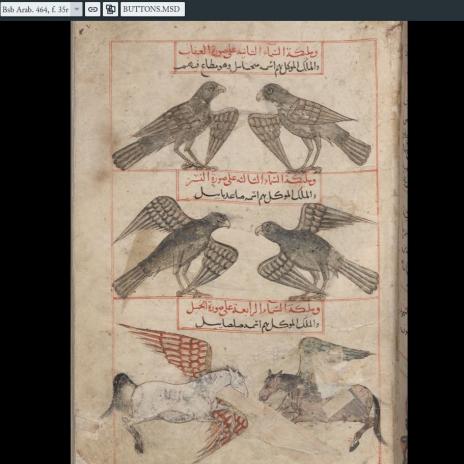
Actividades de divulgación (virtuales o físicas)

- Actividudes de divalgación (virtuales o p
 - vídeo animación Skylitzes
- impacto social (desigualdad) y ecológico (consumo de agua)
- dependencia de infraestructura tecnológica (privada y extranjera)
- bajo (o falta de) apoyo a nivel técnico por parte de las instituciones promotoras del uso de la IA (esp. personal técnico, oferta de formación)

 \rightarrow Estándares web (w3c)

- almacenamiento y preservación a corto/medio/largo plazo
- seguridad de los datos (de investigación y personales)

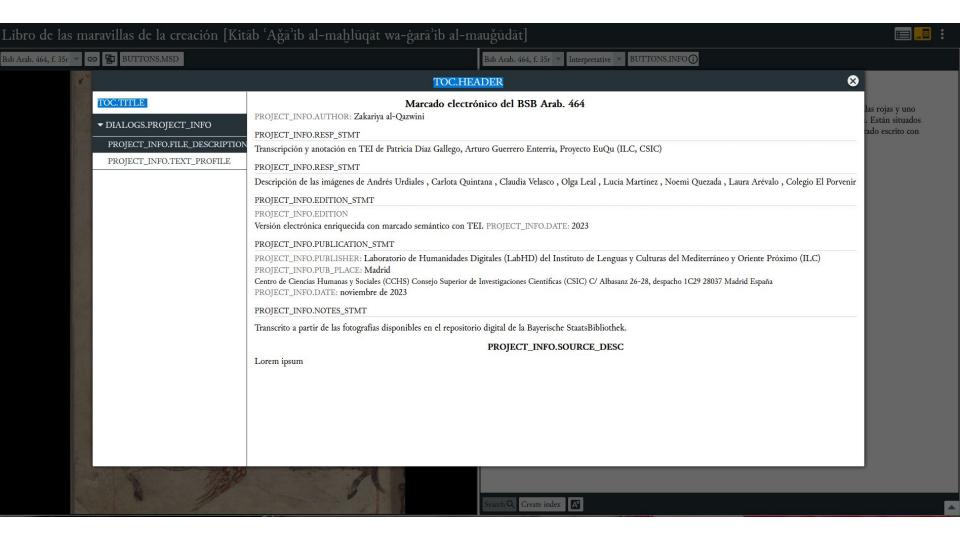
- Nulla aesthetica sine ethica, nulla intelligentia artificialis sine ethica



BUTTONS.INFO() Bsb Arab. 464, f. 35r Interpretative *

pegasos enfrentados

En la parte de abajo del folio. Ocupa un tercio del folio. Salen dos pegasos enfrentados, uno blanco de alas rojas y uno marrón con alas verdes. El marrón no está muy bien conservado, porque no aparece la imagen completa. Están situados dentro de un cuadrado rojo, con un título en rojo y un subtítulo en negro en la parte superior del cuadrado escrito con alfabeto árabe.



De animalium proprietate



Harvard Ms Typ 222, f. 7v Tinterpretative





El sacrificio materno

tEn la primera imágen se observa a un cisne en su nido sacrificandose para alimentar a sus crias hambrientas.En la segunda imágen se ve a un cisne adulto que deducimos que es una de esas crias en un lago. Su aspecto en compartación con el de arriba muestra un gran desarrollo por el sacrificio materno.

▼ BUTTONS.INFO(i)

- título de la imagen
- texto
- <pb n="Harvard Ms Typ 222, f. 7v" facs="https://ids.lib.harvard.edu/ids/iiif/422128664/full/pct:50/0/default.jpg" xml:id="Dik12230_f007v"/>
- El sacrificio materno
- <tEn la primera imágen se observa a un cisne en su nido sacrificandose para alimentar a sus crias hambrientas. En la segunda imágen se ve a un cisne adulto que deducimos que es</p> una de esas crias en un lago. Su aspecto en compartación con el de arriba muestra un gran desarrollo por el sacrificio materno.
- <pb n="Harvard Ms Typ 222, f. 8r" facs="https://ids.lib.harvard.edu/ids/iiif/422128665/full/pct:50/0/default.jpg" xml:id="Dik12230_f008r"/>
- título de la imagen

```
TOC.TITLE
▼ DIALOGS.PROJECT_INFO
```

PROJECT_INFO.FILE_DESCRIPTION

PROJECT_INFO.TEXT_PROFILE

Marcado electrónico del Harvad Ms Typ 222

PROJECT INFO.AUTHOR: Manuel Philes

PROJECT INFO.RESP STMT

Transcripción y anotación en TEI de Carmen García Bueno, ILC, CSIC

<TEI xmlns="http://www.tei-c.org/ns/1.0">

<edition n="1">

PROJECT_INFO.RESP_STMT

Descripción de las imágenes de Ethan Molinero Estevez, Adriano Geitz Sampedro, Alejandra Pérez Pérez, Luna Sánchez Vacas, Rubén Morillo Sánchez, Ignacio Isaac Pérez Jiménez, Samuel Myeong So, Walter Enrique Delgado Pilatasig, Colegio El Porvenir

PROJECT_INFO PROJECT_INFO PROJECT_INFO Centro de Cienci PROJECT_INFO PROJECT INFO Transcrito a par Lorem ipsum

PROJECT INFO PROJECT_INFO Versión electrón

<title>Marcado electrónico del Harvad Ms Typ 222</title> <author xml:id="MPh">Manuel Philes</author> <resp>Transcripción y anotación en TEI de </resp> <name xml:id="CGB">Carmen García Bueno</name> <orgName>ILC, CSIC</orgName> <resp>Descripción de las imágenes de </resp> <name xml:id="EME">Ethan Molinero Estevez <name xml:id="AGS">Adriano Geitz Sampedro <name xml:id="APP">Alejandra Pérez Pérez </name><name xml:id="LSV">Luna Sánchez Vacas

> </name><name xml:id="RMS">Rubén Morillo Sánchez <name xml:id="IIPJ">Ignacio Isaac Pérez Jiménez</name>

<name xml:id="WEDP">Walter Enrique Delgado Pilatasig

<title>Versión electrónica enriquecida con marcado semántico con TEI.</title>

<name xml:id="SMS">Samuel Myeong So

<orgName>Colegio El Porvenir</orgName>

<date when="2023">2023</date>

turas del Mediterráneo y Oriente Próximo (ILC)

despacho 1B5 28037 Madrid España

arvard.

5. Divulgación, preservación *et al.* (*et IA*?)

Otros recursos útiles

→ - Bibliografía (<u>ejemplo</u>)

- Colecciones digitales: Omeka (ejemplo)

Mapas: Recogito (ejemplo), líneas de tiempo (<u>Timelinejs</u>)

- Mi trabajo, para todos los públicos

- Principios FAIR + Ciencia Abierta (MICIIN, UNESCO)

- Findable (Encontrable)

- Accessible (Accesible)

Interoperable (Interoperable)

- Reusable (Reutilizable)

Actividades de divulgación (virtuales o físicas)

- Actividudes de divalgación (virtuales o p
 - vídeo animación Skylitzes
- impacto social (desigualdad) y ecológico (consumo de agua)
- dependencia de infraestructura tecnológica (privada y extranjera)
- bajo (o falta de) apoyo a nivel técnico por parte de las instituciones promotoras del uso de la IA (esp. personal técnico, oferta de formación)

 \rightarrow Estándares web (w3c)

- almacenamiento y preservación a corto/medio/largo plazo
- seguridad de los datos (de investigación y personales)

- Nulla aesthetica sine ethica, nulla intelligentia artificialis sine ethica

Créditos y referencias de las imágenes

- imagen del CAAL, tomada de Jones, S. «Reverse Engineering the First Humanities Computing Center». *Digital Humanities Quarterly* 12, n.º 2 (2018). https://www.digitalhumanities.org/dhq/vol/12/2/000380/000380.html.
- imágenes de la sala de griego, tomadas de «DGE: Presentación». Accedido 21 de abril de 2025. http://dge.cchs.csic.es/hist.
- diagrama tomado de van Lit, C. *Three levels of DH*. 01/10/2021. https://digitalorientalist.com/2021/10/01/how-to-level-up-in-dh/
- ff. 28r (detalle) y 55r (detalle), tomados de *Synopsis historiarum… / conscripta a Joanne Scylitze*, Madrid, Biblioteca Nacional, VITR/26/2, http://bdh-rd.bne.es/viewer.vm?id=0000022766&page=1
- imagen del *kathisma* del hipódromo tomada de Calahorra Bartolomé, A. *Keleusate : arquitectura, arte y ceremonia en la gran palacio de Constantinopla*. Tesis doctoral. Universidad Complutense de Madrid, 26/5/2023, p. 875. https://hdl.handle.net/20.500.14352/92817

Gracias por la atención

Carmen García Bueno ILC (CSIC)

https://ilc.csic.es/es/personal/carmen-garcia-bueno carmen.garcia.bueno@cchs.csic.es





















