

Instituto de Historia del CSIC

LaboratoriosIH

Nombre del Laboratorio MEDlab

Departamento de Estudios Medievales, grupo **REDES DE PODER EN LAS SOCIEDADES MEDIEVALES**

🌐 Enlace a la web del Laboratorio: <https://ih.csic.es/es/org-structure/laboratorio-tratamiento-datos-fuentes-medievales-medlab>

🌐 ✉ Correo de contacto: medlab@cchs.csic.es

- **Descripción:**

Laboratorio de Tratamiento de Datos y Fuentes Medievales (MEDlab).

El Laboratorio de Tratamiento de Datos y Fuentes Medievales (MEDlab) es una unidad de servicio del Instituto de Historia (IH) del CSIC vinculada al Grupo de Investigación *Redes de Poder en las Sociedades Medievales*. Su objetivo es preservar, actualizar y expandir los recursos digitales para la investigación medievalista en España. Estos recursos, generados por los proyectos de investigación internacionales y nacionales coordinados por los miembros del grupo, incluyen bases de datos relacionales y georreferenciadas y repositorios de imágenes. La innovación y el potencial de crecimiento se vinculan a los nuevos desarrollos y capacidades en Humanidades Digitales, que proporcionan resultados cuantitativa y cualitativamente muy relevantes.

Sus objetivos son:

1. Garantizar la continuidad y sostenibilidad de los recursos generados por el Grupo de Investigación *Redes de Poder en las Sociedades Medievales*, que abarcan el análisis de fuentes textuales, materiales y artísticas.



2. Generar nuevos productos mediante la agregación inteligente de datos y la aplicación de metodologías de Inteligencia Artificial y Humanidades Digitales.

3. Consolidarse como laboratorio de referencia nacional e internacional, estableciendo alianzas con centros y laboratorios de excelencia en Europa.

MEDlab presta servicio al Grupo de Investigación *Redes de Poder en las Sociedades Medievales*, a otros grupos del IH y ofrece servicios especializados externos (OPIs y Universidades), como informes técnicos, reconstrucciones 3D, desarrollo de ontologías y cartografía histórica.

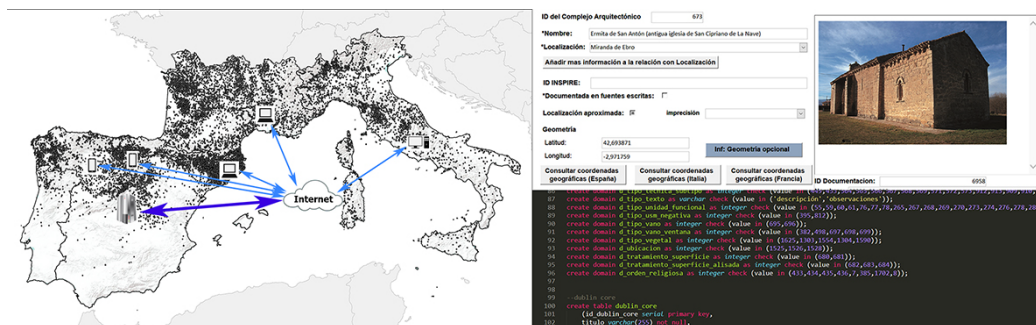
- **Años de funcionamiento:** 4 años
- **Responsable:** Ana Rodríguez
- **Espacios:** Laboratorio de Tratamiento de Datos y Fuentes Medievales (MEDlab)

Planta 2, Módulo E, Despacho 2E28

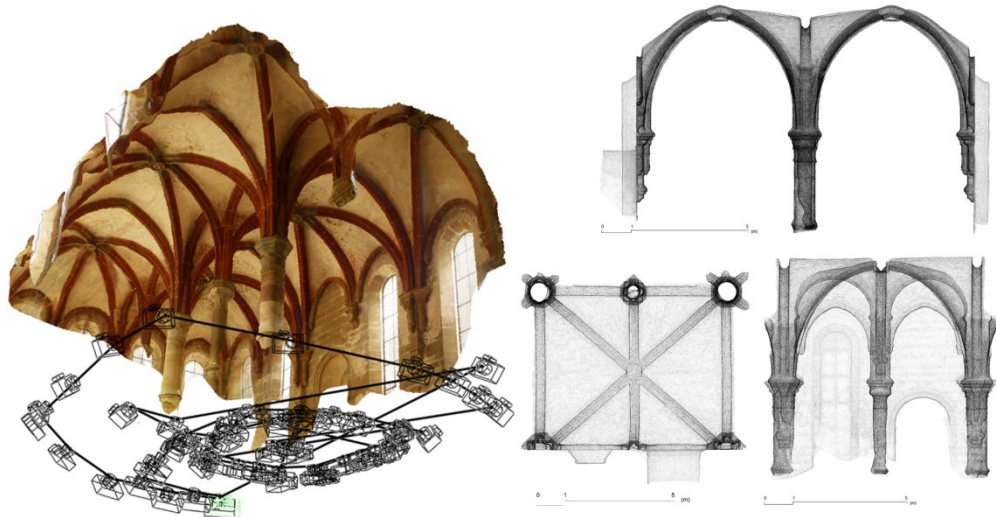
• Uso propio: X

- **Servicios Científico-técnicos:**

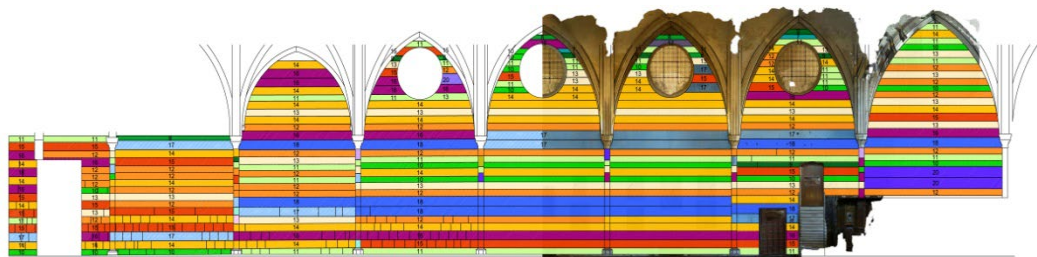
1. Elaboración y tratamiento de conjuntos de datos gráficos procedentes de fuentes medievales.



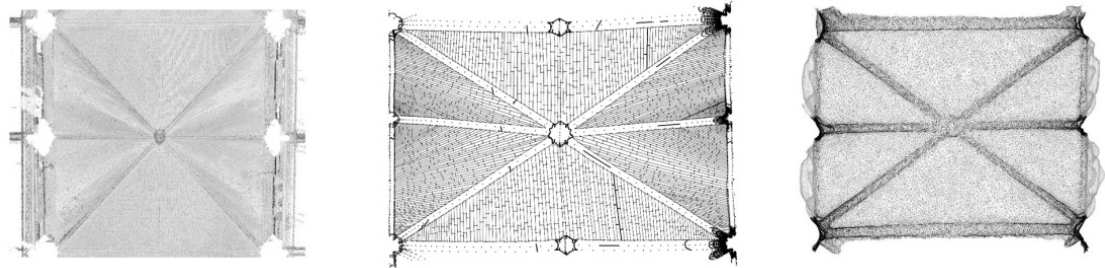
- **Generación de nubes de puntos y modelos tridimensionales de alta precisión a partir de mediciones topográficas, escaneado láser o fotogrametría.**



- **Elaboración de planimetrías, ortofotografías y modelos 3D** derivados de levantamientos topográficos y fotogramétricos.
- **Elaboración de planos constructivos analíticos** de muros, fábricas y otros elementos arquitectónicos para el estudio de los sistemas constructivos.



- **Reconstrucciones tridimensionales de edificios, ciudades, estructuras u objetos desaparecidos**, realizadas a partir de fuentes medievales, evidencias materiales y análisis arquitectónico.
- **Análisis constructivo e interpretación histórica de edificios medievales** a partir de la documentación arquitectónica y material.



- **Análisis e interpretación de referencias a arquitectura y procesos constructivos en fuentes documentales medievales.**

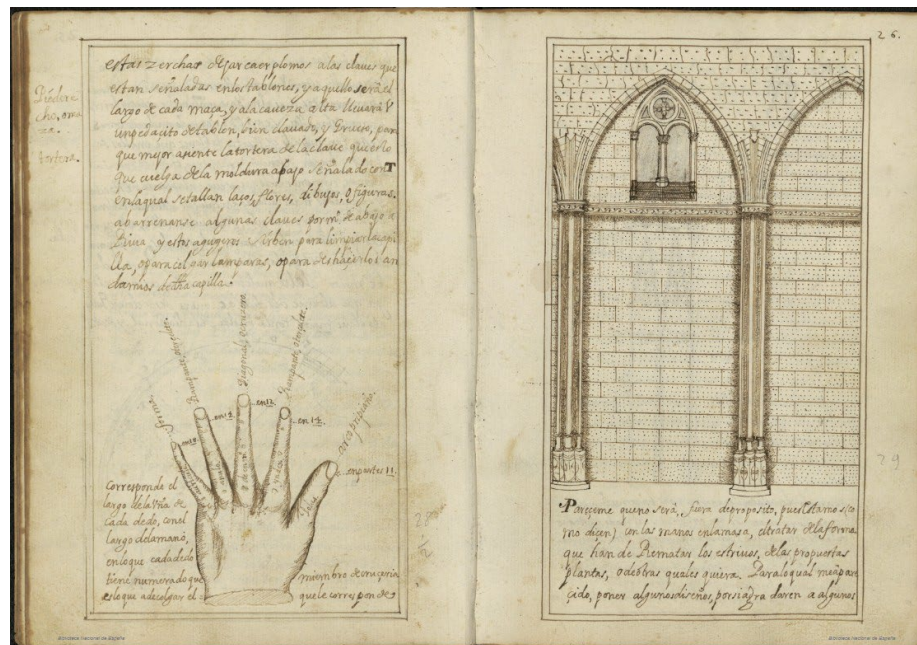
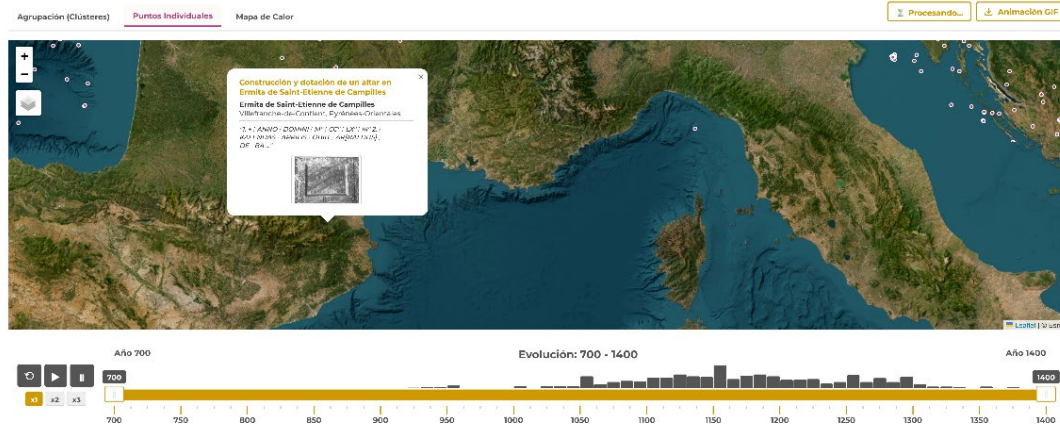


Imagen: Compendio de arquitectura y simetría de los templos. Simón García (fl. 1651-1681), Rodrigo Gil de Hontañón (1500-1577). Biblioteca Nacional de España (CC BY 4.0)

- **Extracción y sistematización de datos de investigación sobre construcción medieval** a partir de informes técnicos, memorias arqueológicas y documentación arquitectónica.
- **Elaboración de conjuntos de datos y bases de información** sobre técnicas constructivas medievales.
- **Análisis de geoespacial y cartografía histórica (SIG)**



- Desarrollo de aplicaciones web.
- Diseño de bases de datos históricas, arqueológicas e iconográficas
- Procesamiento de Lenguaje Natural (PNL) e IA aplicada a textos medievales

```

SimArTI-Searchpy
Arabic Similarity Search Tool for the OpenITI Corpus (SimArTI) > SimArTI-Searchpy > preprocess_arabic_text_for_tfidf > fs
164 def preprocess_arabic_text_for_tfidf(
165     stopwords_ar: List[str],
166     no_numbers: Dict[int, None],
167     nlp: stanza.Pipeline,
168     fs: FarasaStemmer
169 ) -> str:
170     """Arabic preprocessing for similarity: keep Arabic chars, remove numbers, stem, tokenize, remove stopwords/punct."""
171     s = AR_ONLY_RE.sub("", text)
172     s = s.translate(m, numbers)
173     s_stem = fs.stem(s)
174     doc = nlp(s_stem)
175     tokens = [w.text for sent in doc.sentences for w in sent.words]
176     cleaned = [
177         t for t in tokens
178         if t and (t not in stopwords_ar) and (t not in string.punctuation)
179     ]
180     return " ".join(cleaned)
181
182 def split_text_into_fragments(text: str, fragment_size: int) -> List[str]:
183     """Split into fixed-size fragments (characters)."""
184     return [text[i:i + fragment_size] for i in range(0, len(text), fragment_size)]
185
186 # Similarity computation
187 def cosine_matches(
188     query_clean: str,
189     vectorizer: TfidfVectorizer,
190     tfidf_matrix: Any,
191     threshold: float
192 ) -> List[Tuple[int, float]]:
193     """Return [(fragment_index, score), ...] above threshold."""
194     q_vec = vectorizer.transform([query_clean])
195     sims = cosine_similarity(q_vec, tfidf_matrix).flatten()
196     return [(i + 1, float(s)) for i, s in enumerate(sims) if s > threshold]
197
198 # Translation (optional)
199 def translate_text(text: str) -> str:
200     """Translate Arabic->English with fallback. Returns empty string if disabled/unavailable."""
201     if not text:
202         return ""
203
  
```

2. Asesoría, apoyo y coordinación de actividades científico/técnicas

- Asesoramiento y redacción de informes técnicos especializados sobre arquitectura, construcción y materialidad de elementos arquitectónicos históricos.
- Asesoramiento, apoyo técnico y coordinación durante el proyecto



- Asesoramiento, apoyo y coordinación de actividades científicas, difusión y transferencia.



Equipamiento científico-técnico:

Cámaras fotográficas, cámaras de video.

Programas: Rhinoceros, Agisoft Metashape, AutoCAD, Photoshop, Blender, Cloudcompare, MeshLab, Python, PostgreSQL, SQLite, QGIS, ArcGIS Online, Visual Studio Code.