

histoGraph and the CUBRIK project

A collaborative research project supported by the European Commission



histoGraph was developed by the CVCE and its partners as part of the European research project CUBRIK, funded by the European Commission's FP7 programme. CUBRIK incorporates computer science research, the design of human-computation tasks, data visualisation, social engineering and the humanities.

histoGraph offers an effective interface that enables users to access collections of historical sources and discover links among and entities within them. In its current form, histoGraph creates a social graph of people who appear in photos related to the history of European integration. This graph is automatically enriched with relevant sources. To build it, the faces in the photos first need to be identified. This work is shared by humans and machines, and the input from one improves the performance of the other. histoGraph demonstrates new approaches that encourage the public to become involved in humanities research and facilitate exchanges between users.

histoGraph is based on a curated collection of more than 3 000 images which represent the main events and players in the history of European integration. This image collection is managed and hosted by the CVCE. To prepare the photos for the graph, an image indexation pipeline is used to detect individual faces in the photographs. A crowd of 'clickworkers' with no specific training double-checks whether the algorithms have correctly detected all the faces and whether they have missed any. The next step is to identify the faces. An automatic face recognition process is triggered to associate each of the faces with a list of ten possible identities. This list of candidates is then disseminated via Twitter or email to a crowd of experts who vote and comment on their preferred identity.

The image metadata as well, for example people's names and the time and place an image was taken as well as contextual information about associated historical events, can be reviewed by expert users

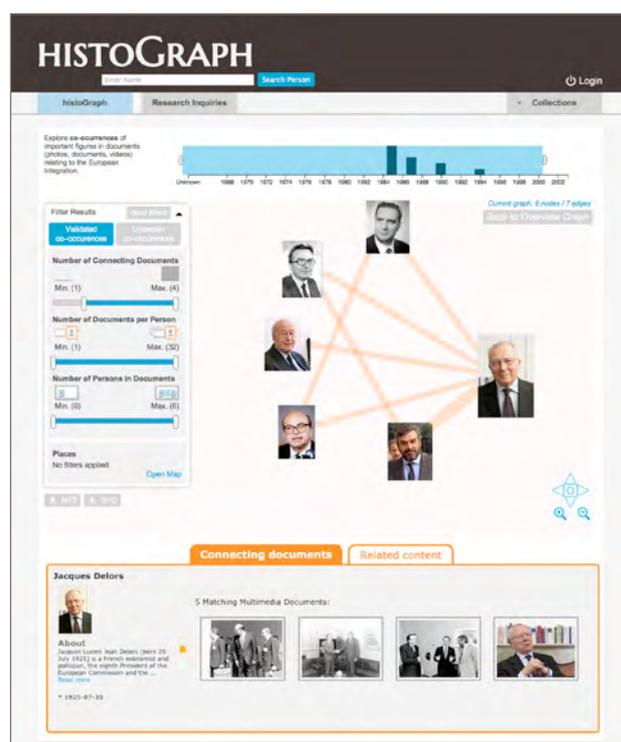
and delegated to a crowd of specialists on the history of European integration for review.

Based on the co-occurrence of people in images, a social graph is created to link individuals with each other. The more times people appear together in different images, the stronger the connections become.

Users can interact with the History of Europe social graph in different ways, for example filtering the graph using a timeline, selecting a network of individual players and exploring which documents connect two people.

Crowd discussion and a new representation of truth

Another challenge for histoGraph and digital humanities in general is the notion of truth. In the humanities the concept of truth is highly complex: people might change careers, their home countries might be renamed or they might choose to go by different names. This means that at different points in



Web interface of the histoGraph application



Knowing the past to build the future

Centre Virtuel de la Connaissance sur l'Europe
Château de Sanem | L-4992 Sanem | Luxembourg
Tel. +352 59 59 20-1 | Fax +352 59 59 20-555 | cvce@cvce.eu

www.cvce.eu



time there might be more than one 'true' answer to the simple question 'Who is this?'. In order to represent the discursive nature of truth in the humanities, histoGraph makes use of a community-driven tool to answer questions. Users are given the opportunity to answer questions themselves and thus benefit from the knowledge within the crowd of experts. The system allows for more than one answer and offers users the possibility to vote, thereby stimulating discussion by allowing them to suggest and defend different opinions.

To learn more about CUBRIK and the modules which have been developed during the project, visit:

<http://www.cvce.eu/digital-humanities>

<http://www.cubrikproject.eu/>

European integration studies in the digital era

The Centre Virtuel de la Connaissance sur l'Europe (CVCE) is a documentation and e-research centre for European integration studies. As a public undertaking under the aegis of the Luxembourg Ministry of Higher Education and Research, the CVCE aims to contribute to a deeper understanding of the historical, legal, economic and political dimensions of European integration. In the pursuit of its mission, it produces knowledge on the European integration process which it makes available to the research and teaching communities and also the wider public. The CVCE is an innovative and committed player in the field of digital humanities. Its interdisciplinary team explores and adopts methods and tools based on digital technologies for its research and knowledge transfer activities in European integration studies.

© European Union

The screenshot shows the histoGraph interface. On the left, a black and white photograph of three men in suits is displayed. Blue bounding boxes are drawn around their faces, with question marks indicating identification points. Below the photo, a list of names is shown: 'Persons' includes 'Lorenzo Natali', 'Pierucci Bondicchi', 'Bruno Kreisky', 'Bettino Craxi', and 'Jacques Delors'. Some names have green checkmarks, while others have question marks. An 'Add a Face' button is also present. To the right of the photo is a metadata form with fields for Title, Date, Event, Place, Photographer, Source, Well Known, and Additional Info. Below the form are fields for Copyright (set to 'Internal use only') and In Collection (set to 'cvce'). On the far right, a 'histoGraph' map shows a network of nodes connected by lines, with a 'Back to Graph' button below it.



The CUBRIK project has received research funding from the European Union, Grant FP7-287704