
CATMA 5.0 Tutorial

Jan Christoph Meister

jan-c-meister@uni-hamburg.de
University of Hamburg, Germany

Evelyn Gius

evelyn.gius@uni-hamburg.de
University of Hamburg, Germany

Jan Horstmann

jan.horstmann@uni-hamburg.de
University of Hamburg, Germany

Janina Jacke

janina.jacke@uni-hamburg.de
University of Hamburg, Germany

Marco Petris

marco.petris@uni-hamburg.de
University of Hamburg, Germany

What is CATMA?

This hands-on tutorial introduces humanists to CATMA (Computer Aided Text Markup and Analysis), a tool developed at the University of Hamburg and currently used by over 60 research projects worldwide. CATMA offers a unique combination of three main features found in no other text analysis tool:

CATMA supports **collaborative annotation and analysis** – a text or text corpus can be investigated individually, but also jointly by a group of students or researchers.

CATMA supports **explorative, non-deterministic practices of text annotation** – a discursive, debate-oriented approach to text annotation based on the research practices of hermeneutic disciplines is the underlying conceptual model.

CATMA integrates **text annotation and text analysis** in a web-based working environment – which makes it possible to combine the identification of textual phenomena with their investigation in a seamless, iterative fashion.

What sets CATMA apart from other digital annotation methods is its ‘undogmatic’ approach: the system does neither prescribe defined annotation schemata or rules, nor does it force the user to apply rigid

yes/no, right/wrong taxonomies to texts (even though it allows for more prescriptive schemata as well). Rather, CATMA’s logic invites users to explore the richness and multi-facetedness of textual phenomena according to their needs: Users can create, expand, and continuously modify their own individual tagsets – so if a text passage invites more than one interpretation, nothing in the system prevents assigning multiple, or even contradictory annotations. Despite all this flexibility, CATMA does not produce idiosyncratic annotations: All markup data can be exported in TEI/XML-format and reused in other contexts.

Since CATMA is a highly intuitive tool, it is also suitable for humanists with little technical knowledge: the GUI allows for a quick kick-off, and CATMA’s query builder (a step-by-step dialogue-based widget) helps users retrieve complex information from texts without having to learn a query language. Another plus on the easy-to-use side is the fact that CATMA’s automated distant-reading functions are continuously enhanced and extended – the current version 5.0 already features a number of automated annotation routines, among others the identification of basic narrative features in texts.

The aim of the tutorial

In our half-day tutorial, we will introduce the core annotation and analysis functionalities of CATMA and show how they can be combined with the annotations provided automatically. Participants will be taken in a step-by-step, hands-on approach through the full cycle of a CATMA-based text investigation:

- From text upload to initial text investigations,
- then to annotation and specification of annotation categories,
- from there to combined text queries that consult the source text and its annotations in combination,

and finally to the visual output of query results.

In a later phase of the tutorial, participants will have the opportunity of testing the tool with regards to their own research interests: They can annotate their own texts or annotate collaboratively a text we will provide. We would also like to engage participants in a critique of CATMA’s design and components as well as a general discussion about requirements for text analysis tools in their fields of interest.

The primary users of CATMA are literary scholars, as well as graduate and undergraduate students of Literary Studies. Nevertheless, this tutorial is likely to be of interest also to:

humanities scholars of ALL fields concerned with text analysis (with and without experience in digital text analysis),

software developers in the humanities interested in non-deterministic text analysis and automated annotation.

Participants need no prior knowledge of digital text annotation and can work with their own laptop computers and their own digital texts. CATMA runs on Laptop or PC (Windows, Unix or MacOS) with a current web browser (MS Explorer or Edge; Firefox, Chrome, Safari) with a mouse or touchpad. Touchscreen navigation is not yet supported (but in the pipeline!). The room in which the workshop takes place should accommodate 25–30 people and provide WLAN and a projector.

Tutorial Instructors

All tutorial instructors come from the developing team of the CATMA project and/or the forTEXT project that is building a platform starting from CATMA. We have been presenting and teaching CATMA on various national and international occasions in the last years.

Evelyn Gius

Evelyn is working in the field of Digital Humanities as a researcher and has been involved in the creation of CATMA from the very beginning. Her research focus is on manual and automated text analysis. For her PhD project in Literary Studies she has explored with CATMA the benefits of applying narratological categories from literary studies to the analysis of narrations of labor conflicts.

Jan Horstmann

Jan uses CATMA as a tool for textual analysis in literature studies and narratology. Currently he is investigating works of Goethe with a combination of distant and digital close reading methods. His focus is to improve the usability of digital tools for people with little or no prior knowledge of computing or programming, i.e. researchers from classical literature studies.

Janina Jacke

Janina Jacke has worked in the heureCLÉA project (2013–2016) that was aimed at developing automated annotation routines for CATMA. Since her research focuses on narratology and theory of interpretation, her

main interest in the DH-context lies in working out the theoretical prerequisites for automated literary annotation.

Jan Christoph Meister

Chris is Professor of German literature with a main research focus in the Digital Humanities. As original inventor of CATMA, he has led several projects concerned with the annotation and visualization of literary data and the development and enhancement of DH-tools.

Marco Petris,

Marco is a computer scientist with a strong affinity for the humanities and has been engaged in the creation of CATMA from the very beginning. As a research developer he is involved in all aspects of the design and implementation of tools for the Digital Humanities.