

Interactive Handwritten Text Recognition and Indexing of Historical Documents: the tranScriptorium Project

Abstract

The Seventh Framework Programme European Project, **tranScriptorium**, aims to develop innovative, efficient and cost-effective solutions for the indexing, search and full transcription of **historical handwritten document images**, using modern, holistic Handwritten Text Recognition (HTR) technology. **tranScriptorium** will turn HTR into a mature technology by capitalizing on interactive-predictive techniques for effective and user-friendly computer-assisted transcription.

Expected users of the HTR technology belong mainly to two groups: a) individual researchers with experience in handwritten documents transcription interested in transcribing specific documents. b) volunteers which collaborate in large transcription projects. The HTR technology will become a support in the digitization of the handwritten materials. The outcomes of the **tranScriptorium** tools will be attached to the published handwritten document images. This includes not only full, correct transcriptions, but also partially correct transcription and other kinds of automatically produced metadata, useful for indexing and searching.

Dr. Alejandro H. Toselli received the M.S. degree in electrical engineering from Universidad Nacional de Tucumán (Argentina) in 1997 and the Ph.D. degree in Computer Science from Universitat Politècnica de València (Spain) in 2004. His current research interest lies in the subject of Computer Assisted and Multimodal Interaction in Pattern Recognition Systems: Handwritten Text Recognition/Transcription Applications. He has participated as an active researcher in several research projects funded by Spanish and European institutions. He has teaching experience in imparting different tutorial courses about the "Handwritten Text Recognition" topic. Alejandro Toselli is also a member of the Spanish Society for Pattern Recognition and Image Analysis (AERFAI) and the International Association for Pattern Recognition (IAPR).