Extracting structured information from unstructured human communication is an ubiquitous task. There is a constant need for this task since computers do not understand our unstructured human communication, and we like to use computers to effectively organize our data. The study of performing this task automatically is known as Information Extraction (IE).

Current approaches to IE can largely be divided into two groups. 1) Rule based systems, which work by extracting information according to a set of pre-defined rules. These systems are flexible, and easy to understand but heuristic in nature. In addition the rules must be manually created and maintained. 2) Token classification systems, that works by classifying tokens, usually words, using machine learning. These are elegant and often superior to rule based systems, but require data labeled at the token level. This data is rarely available for IE tasks and must be explicitly created at great cost.

Inspired by the breakthroughs that end-to-end deep learning has had in several other fields, this thesis investigates end-to-end deep learning for information extraction. End-to-end deep learning works by learning deep neural networks that map directly from the input to the output data naturally consumed and produced in IE tasks. Since it learns from the data that is naturally available, for example as the result of a regular business process, it has the potential to be a widely applicable approach to IE.

The research papers presented in this thesis explore several aspects of end-to-end deep learning for IE. The main contributions are: 1) A novel architecture for end-to-end deep learning for IE, which achieve state-of-the-art results on a large realistic dataset of invoices. 2) A novel end-to-end deep learning method for structured prediction and relational reasoning. 3) A natural and efficient input representation of documents with combined text and image modalities.