Things such as organizations, persons, or locations are ubiquitous in all texts circulating on the internet, particularly in the news, forum posts, and social media. Today, there is more written material than any single person can read through during a typical lifespan. Automatic systems can help us amplify our abilities to find relevant information, where, ideally, a system would learn knowledge from our combined written legacy. Ultimately, this would enable us, one day, to build automatic systems that have reasoning capabilities and can answer any question in any human language. In this work, I explore methods to represent linguistic structures in text, build processing infrastructures, and how they can be combined to process a comprehensive collection of documents. The goal is to extract knowledge from text via things, entities. As text, I focused on encyclopedic resources such as Wikipedia. As knowledge representation, I chose to use graphs, where the entities correspond to graph nodes. To populate such graphs, I created a named entity linker that can find entities in multiple languages such as English, Spanish, and Chinese, and associate them to unique identifiers. In addition, I describe a published state-of-the-art Swedish named entity recognizer that finds mentions of entities in text that I evaluated on the four majority classes in the Stockholm-Umeå Corpus (SUC) 3.0. To collect the text resources needed for the implementation of the algorithms and the training of the machine learning models, I also describe a document representation, Docria, that consists of multiple layers of annotations: A model capable of representing structures found in Wikipedia and beyond. Finally, I describe how to construct processing pipelines for large-scale processing with Wikipedia using Docria.