

Digitizing Named Entities Found Within Letters of Agreement

Stephen S. B. Clarke, David L. Nielsen, Swetha Rajkumar and Krishna M. Kalyanam

NASA Ames Research Center, stephen.s.clarke@nasa.gov

NASA Ames Research Center, david.l.nielsen@nasa.gov

NASA Ames Research Center, swetha.rajkumar@nasa.gov

NASA Ames Research Center, krishna.m.kalyanam@nasa.gov

Letters of Agreement (LOAs) are text-based air traffic control documents that contain procedures and actions agreed upon by the different parties, typically two or more FAA facilities, that are subject to an agreement. The documents contain among other things generic constraints, which are explicit and implicit combinations of procedures that limit a flight's trajectory and affects pilot actions. For example, a controller may be required, to assign a specific altitude to an aircraft crossing the boundary between two airspaces. Although LOA generic constraints directly impact the trajectory of an aircraft, they are not currently available in a digital form that can be used for (or directly ingested into automated) flight planning. Instead, the constraints are manually input into an onboard or ground based system. LOA documents are primarily stored at a controlling facility and the generic constraints are implemented by experienced air traffic controllers and pilots primarily using voice instructions. This increases the workload of the controllers, likelihood of error (e.g., due to noisy communication) and makes it impractical for implementation with unmanned aircraft. Therefore, steps must be taken to make existing constraints machine interpretable to enable e.g., automated handoffs which in turn would reduce controller workload. With recent advances in natural language processing, especially the rise in digitization of text documents (e.g., medical documents) and automated extraction of information therein, it is now possible to extract flight specific constraints from LOAs.

The goal of this work is to digitize *named entities* through a combination of natural language processing tasks: named entity disambiguation, toponym resolution, and numeric parsing to extract general constraint components contained within LOAs, herein referred to as Entity Enhancement (EE). Starting with a small list of named entities (e.g., *ARTCC*, *Tower*, *Altitude* and *Speed*), EE can extract the named entities while simultaneously converting the string-based output into a digital format using an ensemble of processes like rule-based gazetteers and syntactic-lexical patterns. The digital format contains a diverse set of information based on the entity label in question, ranging from standardized facility names to units of measure (e.g., feet) and other numeric information. Upon validating our approach using a truth dataset, we show an overall F1-Score of 0.71 for the extraction process.

Looking beyond entity enhancement, we are also working towards the goal of completely digitizing the general constraints by performing EE and fitting them into a standardized exchange model (XM) such as the Aeronautical Information Exchange Model (AIXM). This will allow for easy distribution and dissemination of LOA constraints to air users, better searchability within documents, and enable ingestion into automated flight planning. Finally, we show a preliminary version of the proposed XM architecture and demonstrate how the model can be populated from the EE output.