

Improving NER Research Workflows with SeqScore

Constantine Lignos, Maya Kruse, and Andrew Rueda

<https://github.com/bltlab/seqscore>

- **SeqScore** is a robust one-stop command line tool for working with NER data
- **Validation:** Validates label encoding (BIO, etc.) and offers options for automatic repair
- **Summarization:** Provides an overview of the types and entities in datasets
- **Conversion:** Converts between different label encodings (BIO to BIOES, etc.)
- **Scoring:** Scores system output with configurable handling of invalid label transitions
- **Error analysis:** Provides reports on errors to enable analysis