BAB VII

# PENUTUP

Sebelum menutup pembahasan mengenai tugas akhir ini, bab terakhir menjelaskan mengenai poin-poin kesimpulan yang ditemukan saat pelaksanaan pengerjaan tugas akhir. Yang ditemukan selain kesimpulan dalam bab ini adalah saran-saran dari penulis untuk penelitian topik ini maupun topik lainnya yang menyinggung topik ini. Saran ini disediakan agar dapat dikembangkan dan dapat digunakan dikemudian hari. Berikut ini adalah kesimpulan dan saran yang telah ditemukan.

## Kesimpulan

Subbab pertama, cukup jelas dari namanya, akan menjelaskan mengenai kesimpulan-kesimpulan yang ditemukan selama mengerjakan dan menganalisa topik tugas akhir ini yaitu penggunaan Sequence-to-Set Network kepada Nested NER dalam bahasa Indonesia. Kesimpulan ini mencakup seluruh proses pengerjaan tugas akhir dari awal sampai akhir, juga kendala-kendala yang dilewatkan dan kesimpulan dari pengerjaan dan kendala yang diterima. Kesimpulan tersebut meliputi:

Sebelum mencoba mengerti apa itu tugas pengenalan named entity (atau NER) secara mendalam, perlu diketahui terlebih dahulu apa yang dapat disebut sebagai sebuah named entity. Istilah named entity awalnya dianggap memiliki relasi dekat dengan pembahasan mengenai *rigid designators* oleh Kripke.[[1]](#footnote-1) Namun diskusi named entity mengarah kepada rigid designators menjadi terlalu filosofis. Sehingga dalam penelitian dicarikan penjelasan yang lebih jelas dan ringkas. Sampai saat ini belum ada persetujuan yang resmi dari bidang NLP mengenai definisi resmi NER. Tetapi dalam skenario penelitian NER dapat disimpulkan definisi umum untuk tugas NER maupun arti sebuah named entities ini. Oleh pihak CoNLL 2002 sendiri, named entities adalah frase yang mengandung nama oleh seseorang, suatu organisasi, sebuah lokasi, sebuah waktu dan sejumlah kuantitas.[[2]](#footnote-2) Dengan arti yang jelas dan ringkas, tugas pengenalan named entities menjadi persetujuan secara umum definisinya seperti yang telah dinyatakan pihak CoNLL 2002.

## Saran

Dataset yang digunakan pada penelitian yang dirujuk di tugas akhir ini menggunakan beberapa dataset, tetapi sebagai rujukan utama di penelitian utama ini akan fokus pada dataset GENIA[[3]](#footnote-3). Sebagai informasi tambahan, dataset lain yang digunakan oleh peneliti Sequence-to-Set Network adalah ACE 2004[[4]](#footnote-4), ACE 2005[[5]](#footnote-5), KBP 2017[[6]](#footnote-6). Keempat dataset yang digunakan merupakan beberapa dataset standar yang digunakan dalam penelitian NER bahasa Inggris. Berikut penjelasan singkat untuk dataset yang digunakan dalam Sequence-to-Set Network.

1. Kripke, Saul, Identity and Necessity, M.K. Munitz (ed.). Identity and Individuation. New York: New York University Press, (New York, 1971), pp. 135–64 [↑](#footnote-ref-1)
2. Kim Sang, Erik F. Tjong, Introduction to the CoNLL-2002 Shared Task: Language-Independent Named Entity Recognition, COLING-02: The 6th Conference on Natural Language Learning 2002, (2002). [↑](#footnote-ref-2)
3. Kim, J.D., dkk., GENIA corpus—a semantically annotated corpus for bio-textmining,  Bioinformatics (2003), Vol. 19 Suppl. 1 2003, hal. i180–i182 [↑](#footnote-ref-3)
4. Shachi Language Research Search, ACE 2004 Multilingual Training Corpus, http://shachi.org/resources/593, 2017. [↑](#footnote-ref-4)
5. Linguistic Data Consortium, ACE 2005 Multilingual Training Corpus, http://catalog.Idc.upenn.edu/ldc2006t06, 2018. [↑](#footnote-ref-5)
6. National Institute of Standards and Technology, Text Analysis Conference (TAC) 2017, https://tac.nist.gov/2017/index.html, 2017. [↑](#footnote-ref-6)