I would like to know to what extent the research reported in the papers in the proceedings of the 2016 workshop “*Biomedical natural language processing,”* associated with the Association for Computational Linguistics annual meeting in Berlin, say that (a) the system, and (b) the data, are available. If either or both of them is/are mentioned, please give me the quote. If not, please tell me so.

You can find the proceedings at: http://aclanthology.info/volumes/proceedings-of-the-15th-workshop-on-biomedical-natural-language-processing

What I mean by “system:” this refers to either the code (a program as written in a programming language, readable and modifiable by humans) or the program itself. Some ways to recognize that a system is made available: look for an explicit statement, which will frequently (although not always) mention some form of the words “available” or “download.” You might also find the URL for a site that distributes code, such as GitHub or SourceForge. You might find the URL of a scholastic or personal web site. Or, it may be none of these. These are good ways to start.

What I mean by “data:” this refers to all of (a) the data that is used to build a system, and (b) the data that is used to evaluate a system. I am not interested in data on the system outputs, or in the system outputs themselves. Some ways to recognize that data is available: look for the name of a publicly available data set. In these cases, the authors of the paper will typically mention the name of a shared task that is associated with the data set (e.g. BioCreative, CLEF eHealth, CoNLL, TREC, or BioNLP-ST) and/or a publication that is associated with it. Or, they may mention the name of a well-known data set (e.g. GENIA, CRAFT, or OHSUMED). If you don’t find one of these, look for an explicit statement, which will frequently (although not always) mention some form of the words “available” or “download.” You might also find the URL for a site that distributes code, such as GitHub or SourceForge. You might find the URL of a scholastic or personal web site. Or, it may be none of these. These are good ways to start.

Please note that I need this by Sunday evening, or I can’t use it. Thanks!

IDEAL ANSWER:

Cohen et al., Distribution of negation, system: yes, “The system is available at github.negation.org”, data: no

Chomsky et al., Three perspectives on negation, system: no, data: yes, “The data is available upon request”

How to get this wrong:

I need BioNLP 2016, not BioNLP-ST 2016.