CLAIMS

What is claimed is:

1. A computer-implemented method comprising:

determining that at least one natural language processing (NLP) request is associated with at least one surface form, the NLP request being for a corpus, a database comprising preexisting annotations associated with the corpus, the NLP request comprising model enhancements that are traced back to the at least one surface form in a text of the corpus;

performing an index query related to the at least one surface form to generate index query results, the index query results comprising identification of portions of the corpus affected by the NLP request;

determining a scope of the NLP request related to the database based on the index query results, the scope comprising identification of impacted candidate annotations of the preexisting annotations affected by the NLP request;

causing a NLP service to be performed on the corpus according to the scope and the portions, thereby resulting in updates; and

committing the updates to the database associated with the corpus.

2. The computer-implemented method of claim 1, further comprising displaying a preview of the updates prior to committing the updates to the database.

3. The computer-implemented method of claim 1, wherein the corpus comprises the at least one surface form.

4. The computer-implemented method of claim 1, wherein the portions of the corpus affected by the NLP request comprise identification of at least one or more documents from a plurality of documents in the corpus.

5. The computer-implemented method of claim 1, wherein the portions of the corpus affected by the NLP request comprise identification of at least one or more passages in documents in the corpus.

6. The computer-implemented method of claim 1, wherein the updates comprise new annotations.

7. The computer-implemented method of claim 1, wherein the updates comprise new annotations and modifications to the impacted candidate annotations.

8. The computer-implemented method of claim 1, wherein causing the NLP service to be performed on the corpus according to the scope and the portions occurs in a shorter amount of time as compared to causing the NLP service to be performed over an entirety of the corpus.

9. The computer-implemented method of claim 1, wherein the performing the index query and the determining the scope enable unilaterally provisioning computing capabilities for the committing the updates to the database associated with the corpus.

10. A system comprising:

a memory having computer readable instructions; and

one or more processors for executing the computer readable instructions, the computer readable instructions controlling the one or more processors to perform operations comprising:

determining that at least one natural language processing (NLP) request is associated with at least one surface form, the NLP request being for a corpus, a database comprising preexisting annotations associated with the corpus, the NLP request comprising model enhancements that are traced back to the at least one surface form in a text of the corpus;

performing an index query related to the at least one surface form to generate index query results, the index query results comprising identification of portions of the corpus affected by the NLP request;

determining a scope of the NLP request related to the database based on the index query results, the scope comprising identification of candidate impacted annotations of the preexisting annotations affected by the NLP request;

causing a NLP service to be performed on the corpus according to the scope and the portions, thereby resulting in updates; and

committing the updates to the database associated with the corpus.

11. The system of claim 10, further comprising displaying a preview of the updates prior to committing the updates to the database.

12. The system of claim 10, wherein the corpus comprises the at least one surface form.

13. The system of claim 10, wherein the portions of the corpus affected by the NLP request comprise identification of at least one or more documents from a plurality of documents in the corpus.

14. The system of claim 10, wherein the portions of the corpus affected by the NLP request comprise identification of at least one or more passages in documents in the corpus.

15. The system of claim 10, wherein the updates comprise new annotations, modifications to the impacted candidate annotations, or both.

16. The system of claim 10, wherein causing the NLP service to be performed on the corpus according to the scope and the portions occurs in a shorter amount of time as compared to causing the NLP service to be performed over an entirety of the corpus.

17. A computer program product comprising a computer readable storage medium having program instructions embodied therewith, the program instructions executable by a processor to cause the processor to perform operations comprising:

determining that at least one natural language processing (NLP) request is associated with at least one surface form, the NLP request being for a corpus, a database comprising preexisting

annotations associated with the corpus, the NLP request comprising model enhancements that are traced back to the at least one surface form in a text of the corpus;

performing an index query related to the at least one surface form to generate index query results, the index query results comprising identification of portions of the corpus affected by the NLP request;

determining a scope of the NLP request related to the database based on the index query results, the scope comprising identification of candidate impacted annotations of the preexisting annotations affected by the NLP request;

causing a NLP service to be performed on the corpus according to the scope and the portions, thereby resulting in updates; and

committing the updates to the database associated with the corpus.

18. The computer program product of claim 17, further comprising displaying a preview of the updates prior to committing the updates to the database.

19. The computer program product of claim 17, wherein the corpus comprises the at least one surface form.

20. The computer program product of claim 17, wherein the portions of the corpus affected by the NLP request comprise identification of at least one or more documents from a plurality of documents in the corpus.