WHAT IS CLAIMED IS:

1. A method comprising:

obtaining, by a computer, interaction data comprising a plurality of interaction data sets, each interaction data set comprising a plurality of tokens, wherein the plurality of tokens include non-repeatable tokens and repeatable tokens;

for each interaction data set:

encoding, by the computer, the non-repeatable tokens of the plurality of tokens to form non-repeatable token embeddings,

retrieving, by the computer, repeatable token embeddings from memory using the repeatable tokens, wherein the repeatable token embeddings and the non-repeatable token embeddings are in a plurality of first embeddings,

concatenating, by the computer, the plurality of first embeddings with positional encodings to form second embeddings,

generating, by the computer, an interaction data set embedding based on the second embeddings;

determining, by the computer, one or more predicted interaction data set embeddings using interaction data set embeddings comprising the interaction data set embedding for each interaction data set; and

decoding, by the computer, the one or more predicted interaction data set embeddings to predicted tokens.

2. The method of claim 1, wherein the plurality of tokens include partially repeatable tokens.

3. The method of claim 2, wherein the non-repeatable tokens include a timestamp and a security value.

4. The method of claim 2, wherein the repeatable tokens include a primary account number, an issuer identifier, an acquirer identifier, a zip code, a resource provider category code, and a resource provider identifier.

5. The method of claim 2, wherein the partially repeatable tokens include a combination of a resource provider identifier and an amount.

6. The method of claim 1, wherein determining one or more predicted interaction data set embeddings comprises:

determining, by the computer, the one or more predicted interaction data set embeddings using a transformer.

7. The method of claim 1, wherein encoding the non-repeatable tokens comprises:

for each non-repeatable token of the non-repeatable tokens, processing, by the computer, the non-repeatable token using a machine learning model to determine a non-repeatable token embedding corresponding to the non-repeatable token.

8. The method of claim 1, wherein the plurality of tokens include partially repeatable tokens, wherein the method further comprises:

for each interaction data set:

separating, by the computer, each partially repeatable token of the partially repeatable tokens into a repeatable token portion and a non-repeatable token portion;

encoding, by the computer, the non-repeatable token portion for each partially repeatable token to form a non-repeatable token portion embedding for each partially repeatable token using a first machine learning model;

retrieving, by the computer, a repeatable token portion embedding for each repeatable token portion from the memory using the repeatable token portion; and

generating, by the computer, partially repeatable token embeddings for each partially repeatable token based on the non-repeatable token portion embedding and the repeatable token portion embedding using a second machine learning model, wherein the partially repeatable token embeddings are in the plurality of first embeddings.

9. The method of claim 1, wherein generating the interaction data set embedding comprises:

generating, by the computer, the interaction data set embedding based on the second embeddings using one or more transformer blocks.

10. The method of claim 1, wherein decoding the one or more predicted interaction data set embeddings comprises:

for each predicted interaction data set embedding of the one or more predicted interaction data set embeddings, generating, by the computer, intermediate embeddings;

for each predicted interaction data set embedding, determining, by the computer, output embeddings based on the intermediate embeddings using one or more transformer blocks;

for each predicted interaction data set embedding, for each output embedding, concatenating, by the computer, the output embedding with a positional encoding to form a positionally encoded output embedding; and

for each predicted interaction data set embedding, for each positionally encoded output embedding, determining, by the computer, a predicted token based on the positionally encoded output embedding using a machine learning model.

11. The method of claim 1, wherein the predicted tokens include a predicted amount, a predicted resource provider identifier, and a predicted primary account number.

12. A computer comprising:

a processor; and

a computer-readable medium coupled to the processor, the computer-readable medium comprising code executable by the processor for implementing a method according to any of the previous claims.

13. A system comprising:

a database storing interaction data stored by a network processing computer; and

an analysis computer comprising:

a processor; and

a computer-readable medium coupled to the processor, the computer-readable medium comprising code executable by the processor for implementing a method comprising:

obtaining, by a computer, interaction data comprising a plurality of interaction data sets, each interaction data set comprising a plurality of tokens;

for each interaction data set:

encoding, by the computer, the plurality of tokens to obtain a plurality of first embeddings,

concatenating, by the computer, the plurality of first embeddings with positional encodings to form second embeddings,

generating, by the computer, an interaction data set embedding based on the second embeddings;

determining, by the computer, one or more predicted interaction data set embeddings using interaction data set embeddings comprising the interaction data set embedding for each interaction data set; and

decoding, by the computer, the one or more predicted interaction data set embeddings to predicted tokens.

14. The system of claim 13, wherein the plurality of tokens include non-repeatable tokens, repeatable tokens, and partially repeatable tokens, wherein the non-repeatable tokens include a timestamp and a security value, wherein the repeatable tokens include a primary account number, an issuer identifier, an acquirer identifier, a zip code, a resource provider category code, and a resource provider identifier, and wherein the partially repeatable tokens include a combination of the resource provider identifier and an amount.

15. The system of claim 13, wherein the database storing the interaction data comprises one hundred or more entities of interaction data sets.