CLAIMS

1. A computer-implemented method comprising:

determining a dataset comprising a plurality of entries, each entry comprising an entity identifier of a plurality of entity identifiers and one or more associated application identifiers of a plurality of application identifiers, each application identifier identifying an application used by an entity of a respective entity identifier;

determining a set of application pairs from the dataset, each application pair comprising a first application identifier and a second application identifier;

determining a correlation score for each application pair of the set based on an association of the first and second application identifiers of the application pair with the plurality of entity identifiers of the dataset, wherein the correlation score is based on a first number of entity identifiers of the dataset with which the first application identifier of the application pair is associated, a second number of entity identifiers of the dataset with which the second application identifier of the application pair is associated, and a third number of entity identifiers of the dataset with which the first and second application identifiers of the application pair are both associated;

responsive to the correlation score for an application pair exceeding a threshold correlation value:

determining the first and second application identifiers of the application pair as being correlated;

determining a correlation write request comprising the first and second applications identifiers of the application pair; and

providing the correlation write request to a write interface of an app recommender graph interface module for writing to respective tables of an app recommender graph structure of an app recommender database for storing the first and second applications identifiers of the application pair as a correlated pair;

receiving an application recommendation request comprising a candidate application identifier of one or more of the correlated pairs stored in the app recommender database; and

responsive to the application recommendation request, providing, by the app recommender database, one or more suggested application identifiers of respective suggested applications, wherein each of the one or more suggested application identifiers forms a respective correlated pair with the candidate application identifier;

wherein providing the correlation write request to the write interface of the app recommender graph interface module comprises:

determining a plurality of correlation write requests to be written to the app recommender graph structure;

collating the plurality of correlation write requests as a batch write request; and

providing the batch write request to the write interface of the app recommender graph interface module.

2. The method of claim 1, wherein the correlation score is indicative of a frequency of a first and second application of the application pair both being associated with a same organization identifier of the organization identifiers of the dataset relative to the frequency of the first application and the second application being independently associated with the organization identifiers of the dataset.

3. The method of claim 1 or claim 2, wherein determining the correlation score comprises determining pointwise mutual information (PMI) of the application pairs.

4. The method of any one of claims 1 to 3, wherein each entry of the dataset comprises a timestamp indicative of a freshness of a relationship between the entity identifier and the application identifier, and wherein the method further comprises:

prior to determining the set of application pairs from the dataset and responsive to the timestamp being greater than a threshold, determining the entry as stale and excluding it from the dataset.

5. The method of any one of the preceding claims, further comprising:

determining a usage relationship between each of the plurality of application identifiers and the respective entity identifier of the dataset;

for each entity identifier and application identifier pair:

determining a usage write request comprising the entity identifier and application identifier of the pair; and

providing the usage write request to a write interface of a graph interface module for writing to respective tables of a graph structure.

6. The method of claim 5, wherein each entry of the dataset comprises a timestamp indicative of a freshness of a relationship between the entity identifier and the application identifier and where the method further comprises:

prior to determining the usage write request and responsive to the timestamp being greater than a threshold, determining the entry as stale and excluding it from the dataset.

7. The method of any one of the preceding claims, wherein determining the dataset comprises:

for each entry of the dataset, determining connection data between the entity identifier and each application identifier of the one or more associated application identifiers, wherein the connection data is indicative of a connection request being made by the entity of the entity identifier to an application service provider providing the application of the application identifier.

8. The method of claim 7, wherein responsive to determining connection data between an entity identifier of the plurality of entity identifiers and a respective application identifier, inferring that the entity of the entity identifier has used the application of the application identifier.

9. A system comprising:

one or more processors; and

memory comprising computer executable instructions, which when executed by the one or more processors, cause the system to:

determine a dataset comprising a plurality of entries, each entry comprising an entity identifier of a plurality of entity identifiers and one or more associated application identifiers of a plurality of application identifiers, each application identifier identifying an application used by an entity of a respective entity identifier;

determine a set of application pairs from the dataset, each application pair comprising a first application identifier and a second application identifier;

determine a correlation score for each application pair of the set based on an association of the first and second application identifiers of the application pair with the plurality of entity identifiers of the dataset, wherein the correlation score is based on a first number of entity identifiers of the dataset with which the first application identifier of the application pair is associated, a second number of entity identifiers of the dataset

with which the second application identifier of the application pair is associated, and a third number of entity identifiers of the dataset with which the first and second application identifiers of the application pair are both associated;

responsive to the correlation score for the application pair exceeding a threshold correlation value:

determine the first and second application identifiers of the application pair as being correlated;

determine a correlation write request comprising the first and second applications identifiers of the application pair; and

provide the correlation write request to a write interface of an app recommender graph interface module for writing to respective tables of an app recommender graph structure of an app recommender database for storing the first and second applications identifiers of the application pair as a correlated pair;

receive an application recommendation request comprising a candidate application identifier of one or more of the correlated pairs stored in the app recommender database; and

responsive to the application recommendation request, provide, by the app recommender database, one or more suggested application identifiers of respective suggested applications, wherein each of the one or more suggested application identifiers forms a respective correlated pair with the candidate application identifier;

wherein providing the correlation write request to the write interface of the app recommender graph interface module comprises causing the system to:

determine a plurality of correlation write requests to be written to the app recommender graph structure;

collate the plurality of correlation write requests as a batch write request; and

provide the batch write request to the write interface of the app recommender graph interface module.

10. The system of claim 9, wherein the correlation score is indicative of a frequency of a first and second application of the application pair both being associated with a same organization identifier of the organization identifiers of the dataset relative to the frequency of the first application and the second application being independently associated with the organization identifiers of the dataset.

11. The system of claim 9 or claim 10, wherein determining the correlation score comprises determining pointwise mutual information (PMI) of the application pairs.

12. The system of any one of claims 9 to 11, wherein each entry of the dataset comprises a timestamp indicative of a freshness of a relationship between the entity identifier and the application identifier, and wherein the method further comprises:
prior to determining the set of application pairs from the dataset and responsive to the timestamp being greater than a threshold, determining the entry as stale and excluding it from the dataset.

13. The system of any one of claims 9 to 12**,** further configure to:

determine a usage relationship between each of the plurality of application identifiers and the respective entity identifier of the dataset; and

for each entity identifier and application identifier pair:

determine a usage write request comprising the entity identifier and application identifier of the pair; and

provide the usage write request to a write interface of a graph interface module for writing to respective tables of a graph structure.

14. The system of claim 13, wherein each entry of the dataset comprises a timestamp indicative of a freshness of a relationship between the entity identifier and the application identifier and where the system is further configured to:

prior to determining the usage write request and responsive to the timestamp being greater than a threshold, determine the entry as stale and excluding it from the dataset.

15. The system of any one of claims 9 to 14, wherein determining the dataset comprises:

for each entry of the dataset, determining connection data between the entity identifier and each application identifier of the one or more associated application identifiers, wherein the connection data is indicative of a connection request being made by the entity of the entity identifier to an application service provider providing the application of the application identifier.

16. A computer-readable storage medium storing instructions that, when executed by a computer, cause the computer to perform operations including:

determining a dataset comprising a plurality of entries of a plurality of entity identifiers, each entry comprising an entity identifier and one or more associated application identifiers of a plurality of application identifiers, each application identifier identifying an application used by an entity of a respective entity identifier;

determining a set of application pairs from the dataset, each application pair comprising a first application identifier and a second application identifier;

determining a correlation score for each application pair of the set based on an association of the first and second application identifiers of the application pair with the plurality of entity identifiers of the dataset, wherein the correlation score is based on a first number of entity identifiers of the dataset with which the first application identifier of the application pair is associated, a second number of entity identifiers of the dataset with which the second application identifier of the application pair is associated, and a third number of entity identifiers of the dataset with which the first and second application identifiers of the application pair are both associated;

responsive to the correlation score for the application pair exceeding a threshold correlation value:

determining the first and second application identifiers of the application pair as being correlated;

determining a correlation write request comprising the first and second applications identifiers of the application pair; and

providing the correlation write request to a write interface of an app recommender graph interface module for writing to respective tables of an app recommender graph structure of an app recommender database for storing the first and second applications identifiers of the application pair as a correlated pair;

receiving an application recommendation request comprising a candidate application identifier of one or more of the correlated pairs stored in the app recommender database; and

responsive to the application recommendation request, providing, by the app recommender database, one or more suggested application identifiers of respective suggested applications, wherein each of the one or more suggested application identifiers forms a respective correlated pair with the candidate application identifier, wherein providing the correlation write request to the write interface of the app recommender graph interface module comprises:

determining a plurality of correlation write requests to be written to the app recommender graph structure;

collating the plurality of correlation write requests as a batch write request; and

providing the batch write request to the write interface of the app recommender graph interface module.

17. The computer-readable storage medium of claim 16, wherein the correlation score is indicative of a frequency of a first and second application of the application pair both being associated with a same organization identifier of the organization identifiers of the dataset relative to the frequency of the first application and the second application being independently associated with the organization identifiers of the dataset.

18. The computer-readable storage medium of claim 16 or claim 17, wherein determining the correlation score comprises determining pointwise mutual information (PMI) of the application pairs.

19. The computer-readable storage medium of any one of claims 16 to 18, wherein each entry of the dataset comprises a timestamp indicative of a freshness of a relationship between the entity identifier and the application identifier, and wherein the method further comprises: prior to determining the set of application pairs from the dataset and responsive to the timestamp being greater than a threshold, determining the entry as stale and excluding it from the dataset.

20. The computer-readable storage medium of claim 16 to 19**,** further comprising: determining a usage relationship between each of the plurality of application identifiers and the respective entity identifier of the dataset;

for each entity identifier and application identifier pair:

determining a usage write request comprising the entity identifier and application identifier of the pair; and

providing the usage write request to a write interface of a graph interface module for writing to respective tables of a graph structure.