What is claimed is:

1. A method for configuring a first computer executable program, the method including:

through a user interface, receiving information indicative of a source of data and a data target;

displaying, in the user interface, a first prompt for a user to provide input to select a type of a process from among a set of multiple process types displayed in the user interface, each of the multiple process types corresponding to a respective second computer program, each second computer program configurable to receive data records, process data records, and output processed data records;

responsive to receiving a selection of the type of process, displaying, in the user interface, a second prompt for information about one or more characteristics associated with a process of the selected type, wherein the second prompt is based on the selected type of process;

through the user interface, responsive to the second prompt, receiving the information about the one or more characteristics associated with the process of the selected type;

based on the received information indicative of the source of data and the data target, automatically assigning values to respective parameters of the first computer executable program to cause the first computer executable program to, when executed, receive data records from the source of data and output processed data records to the data target; and

based on the selected type of process and the received information about the one or more characteristics of the process, automatically configuring the first computer executable program to reference a second computer program that is configurable to receive data records, process data records, and output processed data records, the configuring including:

automatically identifying the particular second computer program to which the selected type of the process corresponds;

based on the received information about the one or more characteristics of the process, automatically determining values for each of one or more parameters of the particular second computer program, wherein values for parameters of a computer program define a behavior of the computer program when the computer program is executed; and

automatically assigning the determined values to respective parameters of the particular second computer executable program to configure the second computer program with respect to receiving data records, processing data records, and outputting processed data records,

wherein when the first computer executable program is executed, the assignment of the determined values to the respective parameters of the particular second computer program causes the particular second computer program to behave according to the characteristics of the process indicated by the received information.

2. The method of claim 1, in which the first computer executable program includes a first dataflow graph including one or more nodes representing data processing components and one or more links representing flows of data between the data processing components.

3. The method of claim 1 or 2, in which the second computer program includes a second dataflow graph including one or more nodes representing data processing components and one or more links representing flows of data between the data processing components.

4. The method of any of the preceding claims, in which the information indicative of the source of data identifies a type of the source of data.

5. The method of claim 4, in which receiving the information indicative of the source of data includes receiving a selection of the type of the source of data from among a set of one or more types.

6. The method of claim 5, in which the set of one or more types includes one or more of a file, a table, an HDFS dataset, a HIVE table, and a queue.

7. The method of any of the preceding claims, in which the information indicative of the source of data includes information indicative of a location of the source of data.

8. The method of any of the preceding claims, in which the information indicative of the data target identifies a type of the data target.

9. The method of claim 8, in which receiving the information indicative of the data target includes receiving a selection of the type of the data target from among a set of one or more types.

10. The method of claim 9, in which the set of one or more types includes one or more of a file, a table, an HDFS dataset, a HIVE table, and a queue.

11. The method of any of the preceding claims, in which the information indicative of the data target includes information indicative of multiple data targets.

12. The method of any of the preceding claims, in which the information indicative of the data target includes information indicative of a location of the data target.

13. The method of any of the preceding claims, including receiving a selection of each of multiple types of processes.

14. The method of any of the preceding claims, including determining the second prompt based on the type of the process.

15. The method of claim 1, wherein the process types of the set of multiple process types include one or more of a mapping process, a filter process, a deduplication process, an aggregation process, a replicate process, a normalization process, a partition and sort process, a change data capture process, and a join process.

16. The method of any of the preceding claims, including automatically configuring the first computer executable program to reference the second computer program at runtime of the first computer executable program.

17. The method of any of the preceding claims, in which automatically configuring the first computer executable program to reference the second computer program includes configuring the first computer executable program to execute the second computer program during execution of the first computer executable program.

18. A non-transitory computer readable medium storing instructions for causing a computing system to configuring a first computer executable program, the instructions including instructions for causing the computing system to perform the steps of the method of any one of claims 1 to 17.

19. A computing system for configuring a first computer executable program, the computing system including:

one or more processors and a memory, the one or more processors and memory configured to perform the steps of the method of any one of claims 1 to 17.

20. A computing system for configuring a first computer executable program, the computing system including:

means for receiving, through a user interface, information indicative of a source of data and a data target;

means for displaying, in the user interface, a first prompt for a user to provide input to select a type of a process from among a set of multiple process types displayed in the user interface, each of the multiple process types corresponding to a respective second computer program, each second computer program configurable to receive data records, process the data records, and output the processed data records;

responsive to receiving a selection of the type of process, displaying, in the user interface, a second prompt for information about one or more characteristics associated with a process of the selected type, wherein the second prompt is based on the selected type of process;

means for, responsive to the second prompt, receiving, through the user interface, the information about the one or more characteristics associated with the process of the selected type;

means for, based on the received information indicative of the source of data and the data target, automatically assigning values to respective parameters of the first computer executable program to cause the first computer executable program to, when executed, receive data records from the source of data and output processed data records to the data target; and

means for, based on the selected type of process and the received information about the one or more characteristics of the process, automatically configuring the first computer executable program to reference a particular second computer program that is configurable to receive data records, process data records, and output processed data records , the configuring including:

automatically identifying the particular second computer program to which the selected type of process corresponds;

based on the received information about the one or more characteristics of the process, automatically determining values for each of one or more parameters of the particular second computer program, wherein values for parameters of a computer program define a behavior of the computer program when the computer program is executed; and

automatically assigning the determined values to respective parameters of the particular second computer executable program to configure the second computer program with respect to receiving data records, processing data records, and outputting processed data records,

wherein when the first computer executable program is executed, the assignment of the determined values to the respective parameters of the particular second computer program causes the particular second computer program to behave according to the characteristics of the process indicated by the received information.

21. A method for configuring a first computer executable program, the method including:

receiving, through a user interface, a selection of a type of a process from among a set of multiple process types displayed in the user interface, each of the multiple process types corresponding to a respective second computer program, each second computer program configurable to receive data records, process data records, and output processed data records;

responsive to receiving the selection of the type of process, displaying, in the user interface, a prompt for information about one or more characteristics associated with a process of the selected type, wherein the prompt is based on the selected type of process;

through the user interface, responsive to the prompt, receiving the information about the one or more characteristics associated with the process of the selected type;

automatically configuring the first computer executable program to reference a particular second computer program that is configurable to receive data records, process data records, and output processed data records, the configuring including:

automatically identifying the particular second computer program to which the selected type of process corresponds;

based on the received information about the one or more characteristics of the process, automatically determining values for each of one or more parameters of the particular second computer program, wherein values for parameters of a computer program define a behavior of the computer program when the computer program is executed; and

automatically assigning the determined values to respective parameters of the particular second computer program to configure the second computer program with respect to receiving data records, processing data records, and outputting processed data records,

wherein when the first computer executable program is executed, the assignment of the determined values to the respective parameters of the particular second computer program causes the particular second computer program to behave according to the characteristics of the process indicated by the received information.

22. The method of claim 21, in which the first computer executable program includes a first dataflow graph including one or more nodes representing data processing components and one or more links representing flows of data between the data processing components.

23. The method of claim 21 or 22, in which the second computer program includes a second dataflow graph including one or more nodes representing data processing components and one or more links representing flows of data between the data processing components

24. The method of any of claims 21 to 23, including determining the prompt based on the selected type of process.

25. The method of any of claims 21 to 24, wherein the process types of the set of multiple process types include one or more of a mapping process, a filter process, a deduplication process, an aggregation process, a replicate process, a normalization process, a partition and sort process, a change data capture process, or a join process.

26. The method of any of claims 21 to 25, comprising automatically configuring the first computer executable program to reference the particular second computer program at runtime of the first computer program.

27. The method of any of claims 21 to 26, in which automatically configuring the first computer executable program to reference the particular second computer program includes configuring the first computer executable program to execute the second computer program during execution of the first computer executable program.

28. A non-transitory computer readable medium storing instructions for causing a computing system to configuring a first computer executable program, the instructions including instructions for causing the computing system to perform the steps of the method of any one of claims 21 to 27.

29. A computing system for configuring a first computer executable program, the computing system including one or more processors and a memory, the one or more processors and memory configured to perform the steps of the method of any one of claims 21 to 27.