

CLAIMS

1. A method including:

analyzing, by one or more processors, a first version of a computer program, the analyzing including identifying a first process included in the first version of the computer program, the first process corresponding to and configured to perform a first operation on data of a first format;

wherein the first process is defined by a first specification, wherein the first specification identifies the first format of data;

defining, based on the analysis, a second specification based on the first specification, wherein the second specification identifies a second format of data different from the first format;

wherein the second specification defines one or more second processes that are configured to perform a second operation on data of the second format, wherein the second operation is based on the first operation; and

by one or more processors, generating a second version of at least a portion of the computer program based on the first version of the computer program and using the second specification, including:

omitting the first process from the second version, and

including, in the second version of the at least portion of the computer program, the one or more second processes defined by the second specification.
2. The method of claim 1, in which identifying the first process includes identifying a process in which the first operation is dependent on a format of the data.
3. The method of claim 1, in which identifying the first process includes identifying a process that is unable to perform the first operation on data of the second format.

4. The method of claim 1, including determining a format of data to be processed by the first process.
5. The method of claim 4, in which identifying the first process includes identifying a process that is unable to perform the first operation on data having the format of the data to be processed by the first process.
6. The method of claim 1, in which identifying the first process includes identifying a first data processing element of the computer program, the first data processing element configured to execute the first process.
7. The method of claim 6, in which including the one or more second processes in the computer program includes including one or more second data processing elements in the second version of the at least portion of the computer program, the second data processing element configured to execute the one or more second processes.
8. The method of claim 1, in which the first format comprises a data type.
9. The method of claim 1, in which the first format includes a size of a data element.
10. The method of claim 1, in which the first process is configured to perform the first operation on data records of a first record format and in which the one or more second processes are configured to perform the second operation on data records of a second record format.
11. The method of claim 10, in which the first record format comprises a name of a field in the records.
12. The method of claim 1, including presenting, in a user interface, an identifier of the first operation.

13. The method of claim 1, in which generating the second version of at least a portion of the computer program includes generating a copy of the portion of the computer program.
14. The method of claim 13, including modifying the copy of the portion of the computer program to omit the first process and to include the one or more second processes.
15. The method of claim 1, including executing the second version of the computer program.
16. The method of claim 1, wherein the second specification is an overlay specification.
17. The method of claim 16, in which generating the second version of the computer program includes generating the second version based on the first version of the computer program and using the overlay specification.
18. The method of claim 16, in which the overlay specification identifies one or more of a process upstream of the first process and a process downstream of the first process.
19. The method of claim 16, including identifying the first process based on an analysis of executable code defining the first process.
20. The method of claim 1, in which the computer program comprises a graph.
21. The method of claim 20, in which the first process is an executable process represented by a first component of the graph and in which the one or more second processes are executable processes represented by one or more second components of the graph.
22. The method of claim 21, in which the one or more second components are configured to receive data records from an upstream component of the graph.

23. The method of claim 21, in which the one or more second components are configured to provide data records to a downstream component of the graph.
24. A system including one or more processors coupled to a memory, the one or more processors and memory configured to perform the steps in the method of any one of claims 1 to 23.
25. A non-transitory computer-readable medium storing instructions for causing a computing system to perform the steps in the method of any one of claims 1 to 23.
26. A system including:
- means for analyzing, by a processor, a first version of a computer program, the analyzing including identifying a first process included in the first version of the computer program, the first process corresponding to and configured to perform a first operation on data of a first format;
 - wherein the first process is defined by a first specification, wherein the first specification identifies the first format of data;
 - means for defining, based on the analysis, a second specification based on the first specification, wherein the second specification identifies a second format of data different from the first format;
 - wherein the second specification defines one or more second processes that are configured to perform a second operation on data of the second format, wherein the second operation is based on the first operation; and
 - means for generating, by a processor, a second version of at least a portion of the computer program based on the first version of the computer program and using the second specification, the generating including:
 - omitting the first process from the second version, and
 - including, in the second version of the at least portion of the computer program, the one or more second processes defined by the second specification.