

What is claimed is:

1. A method including:
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 configured to perform a first operation on data of a first format; and

by a processor, generating a second version of at least a portion of the computer program, including omitting the first process and including in the second version of the at least portion of computer program one or more second processes configured to perform a second operation on data of a second format different from the first format, wherein the second operation is based on the first operation.
2. The method of claim 1, in which identifying a first process includes identifying a first process in which the first operation is dependent on the format of the data.
3. The method of claim 1 or 2, in which identifying a first process includes identifying a first process that is unable to perform the first operation on data of the second format.
4. The method of any of the preceding claims, including determining a format of data to be processed by the first process.
5. The method of claim 4, in which identifying a first process includes identifying a first 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 any of the preceding claims, in which identifying a 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 any of the preceding claims, in which the first format comprises a data type.
9. The method of any of the preceding claims, in which the first format includes a size of a data element.
10. The method of any of the preceding claims, 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 any of the preceding claims, including presenting, in a user interface, an identifier of the first set of one or more operations.
13. The method of any of the preceding claims, 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 any of the preceding claims, including executing the second version of the computer program.

16. The method of any of the preceding claims, in which the one or more second processes are defined by 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 the overlay specification.
18. The method of claim 16 or 17, 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 any of claims 16 to 18, 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 any of the preceding claims, 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 or 22, 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:
 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 configured to perform a first operation on data having a first format; and

means for generating, by a processor, a second version of at least a portion of the computer program, including omitting the first process and including in the second version of the at least portion of the computer program one or more second processes configured to perform a second operation on data having a second format different from the first format, wherein the second operation is based on the first operation.

25. A system including:

a processor coupled to a memory, the processor and memory configured to:

analyze 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 configured to perform a first operation on data having a first format; and

generate a second version of at least a portion of the computer program, including omitting the first process and including in the second version of the at least portion of the computer program one or more second processes configured to perform a second operation on data having a second format different from the first format, wherein the second operation is based on the first operation.

26. A non-transitory computer-readable medium storing instructions for causing a computing system to:

analyze 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 configured to perform a first operation on data having a first format; and

generate a second version of at least a portion of the computer program, including omitting the first process and including in the second version of the at least portion of the computer program one or more second processes configured to perform a second operation on data having a second format different from the first format, wherein the second operation is based on the first operation.