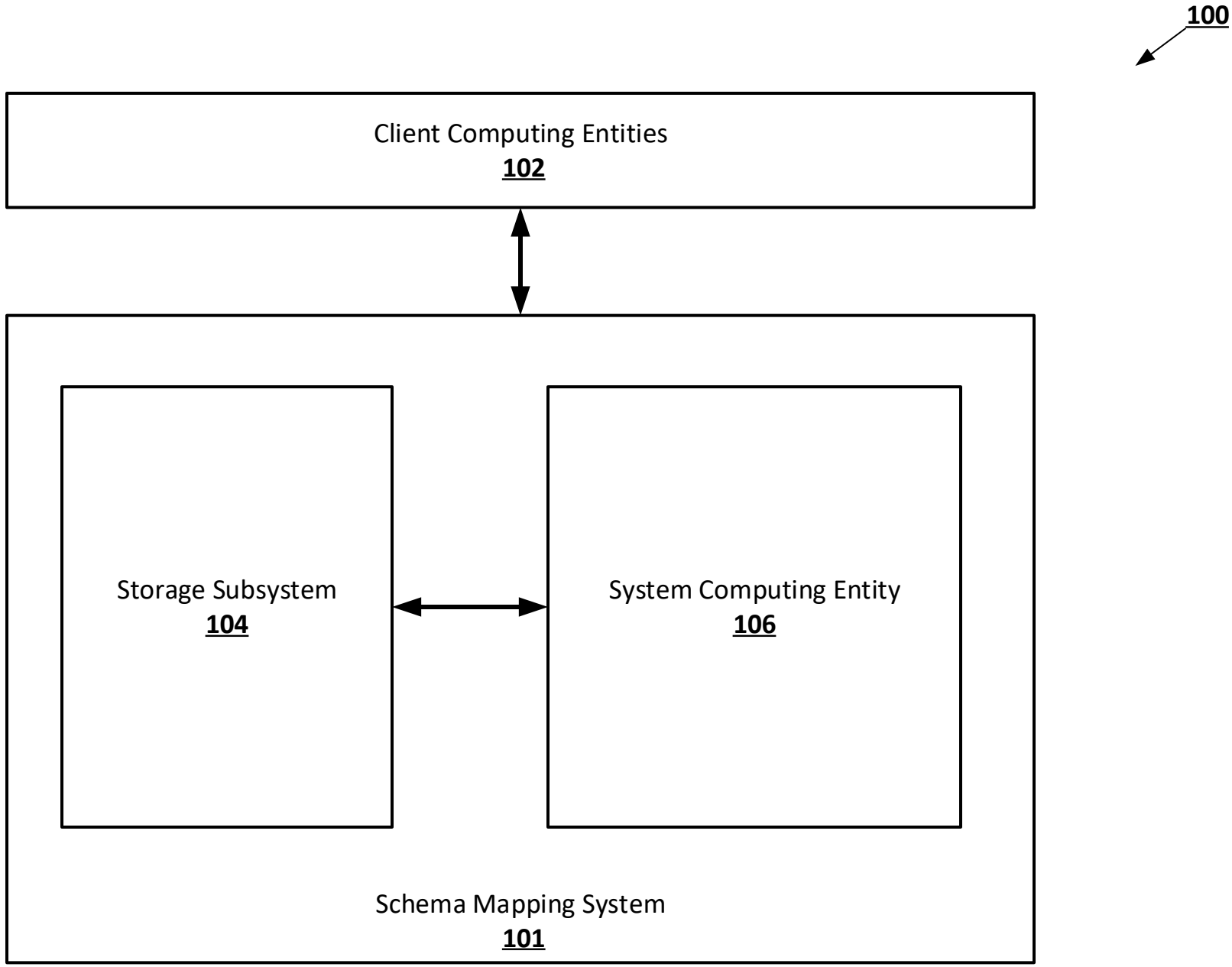

100
Client Computing Entities
102
Storage Subsystem
104
System Computing Entity
106
Schema Mapping System
101

FIG. 1

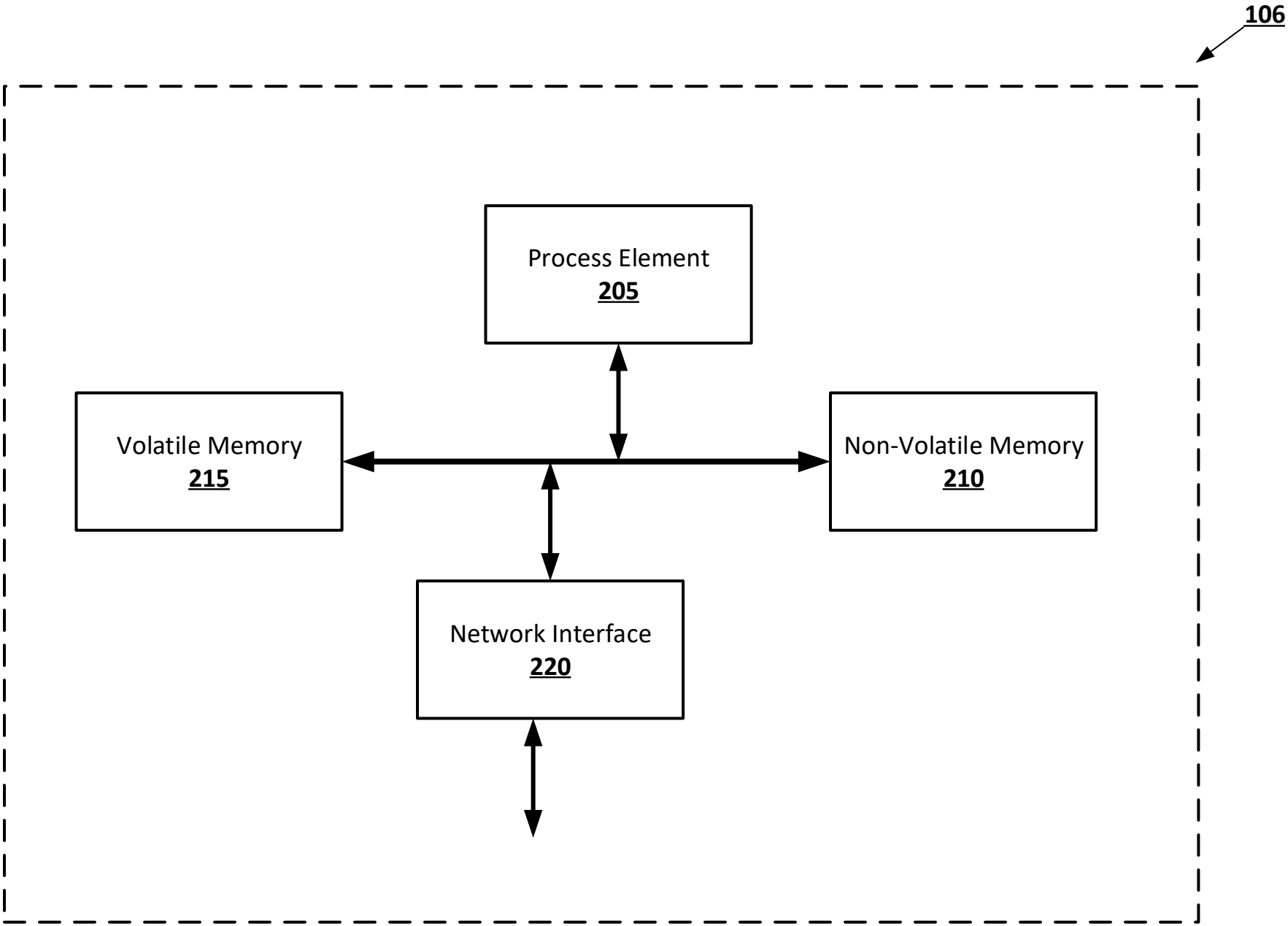

106
Process Element
205
Volatile Memory
215
Non-Volatile Memory
210
Network Interface
220

FIG. 2

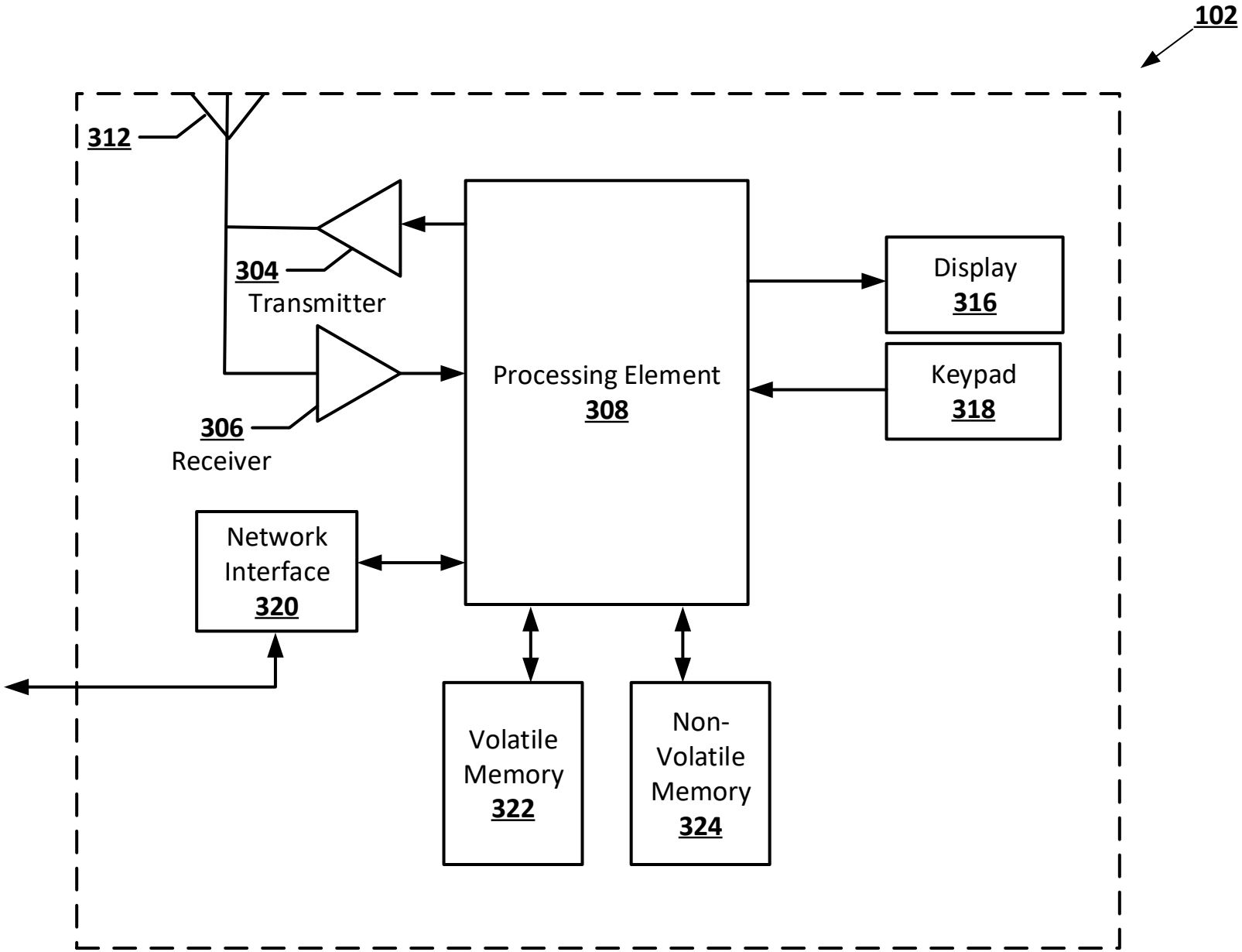

102
312
304
Transmitter
306
Receiver
Processing Element
308
Display
316
Keypad
318
Network
Interface
320
Volatile
Memory
322
Non-
Volatile
Memory
324

FIG. 3

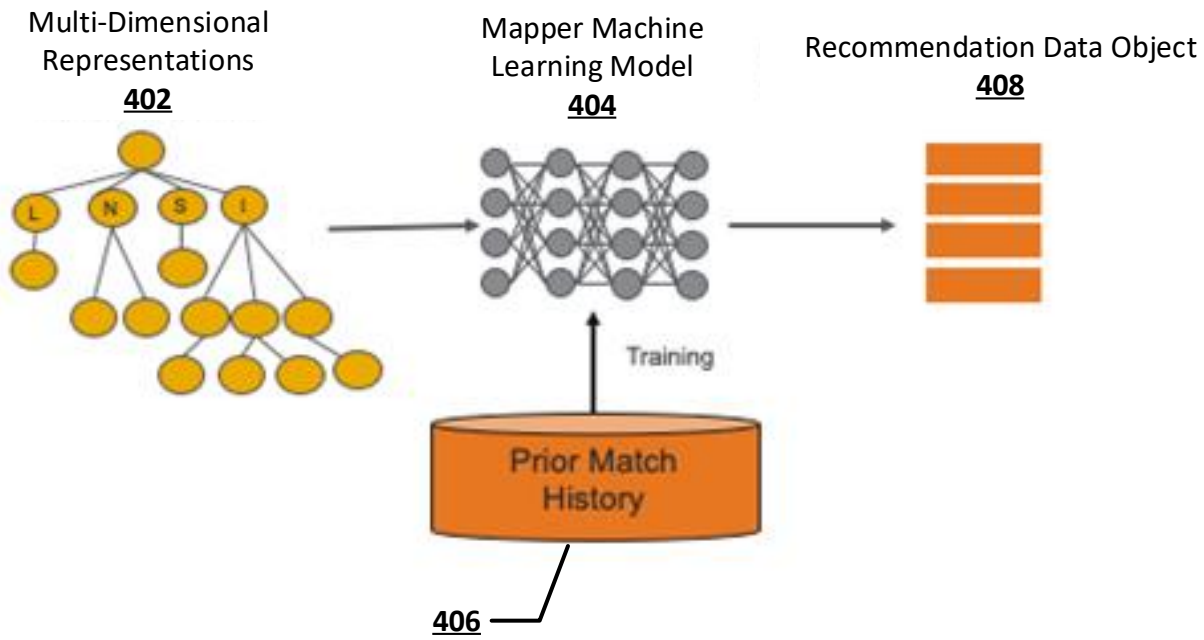

Multi-Dimensional Representations
402
Mapper Machine Learning Model
404
Recommendation Data Object
408
L
N
S
I
Training
Prior Match History
406

FIG. 4

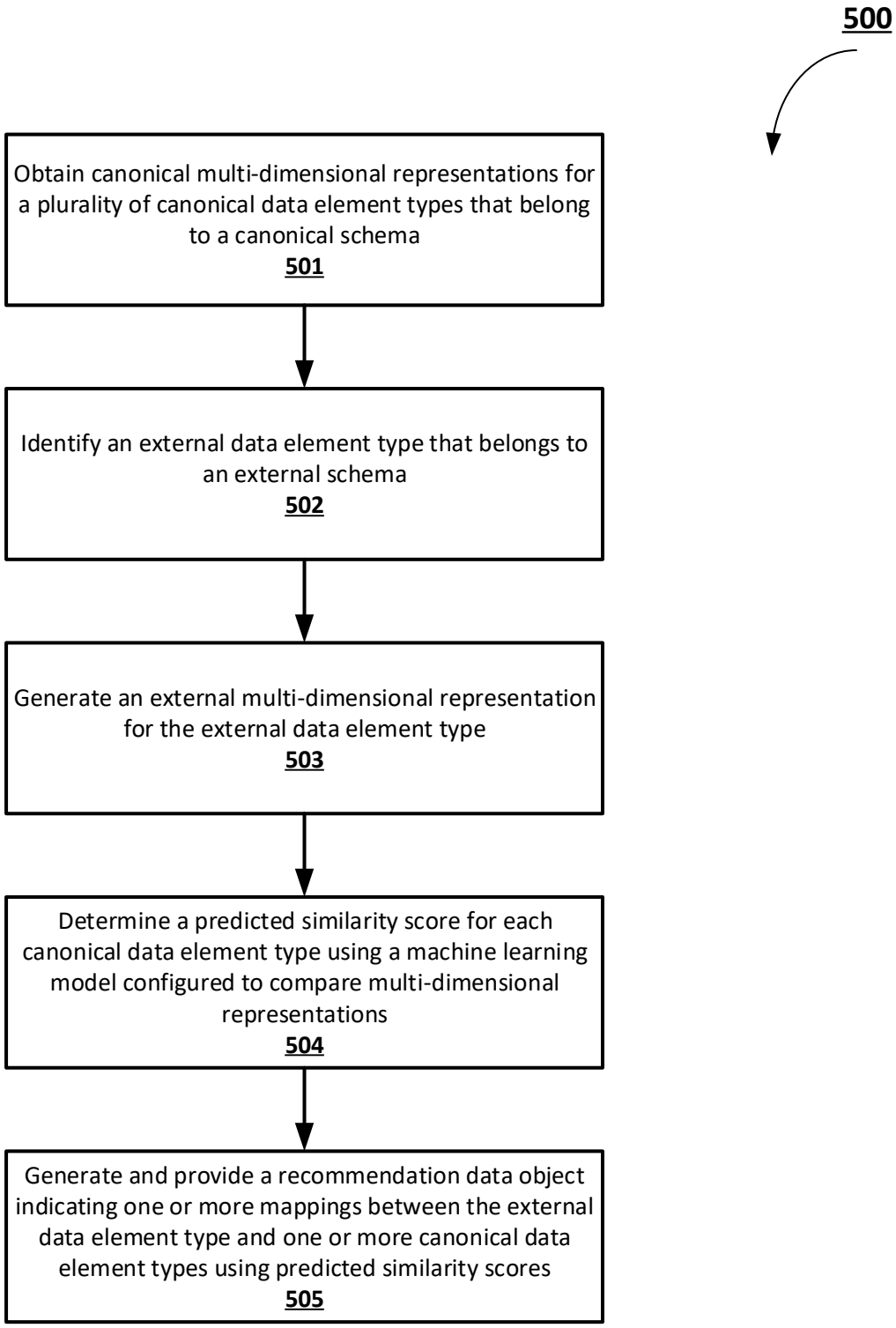

500
Obtain canonical multi-dimensional representations for a plurality of canonical data element types that belong to a canonical schema
501
Identify an external data element type that belongs to an external schema
502
Generate an external multi-dimensional representation for the external data element type
503
Determine a predicted similarity score for each canonical data element type using a machine learning model configured to compare multi-dimensional representations
504
Generate and provide a recommendation data object indicating one or more mappings between the external data element type and one or more canonical data element types using predicted similarity scores
505

FIG. 5

External Data Element Type **600**

602 **DataElementName**	IndividualFirstName
606 **DataElementUri**	http://marklogic.com/metadata-registry/data-element/individual-first-name.xml
DataElementObjectClassCode	Individual
DataElementAbstractIndicator	false
DataElementPropertyTermName	First
DataElementRepresentationTerm	Name
604 **DataElementDefinitionText**	An Individual's First Name or Given Name
SubjectAreas	**SubjectAreaCode** demographic
DataStewards	**PrimaryStewardId** admin **AlternateStewardId** dmccreary
Approvals	**Approval** → **ApprovalTypeCode**, **ApprovdeByUserId**, **ApprovalDateTime**
DataElementStatistics	**xml-schema-type** xs:string

FIG. 6

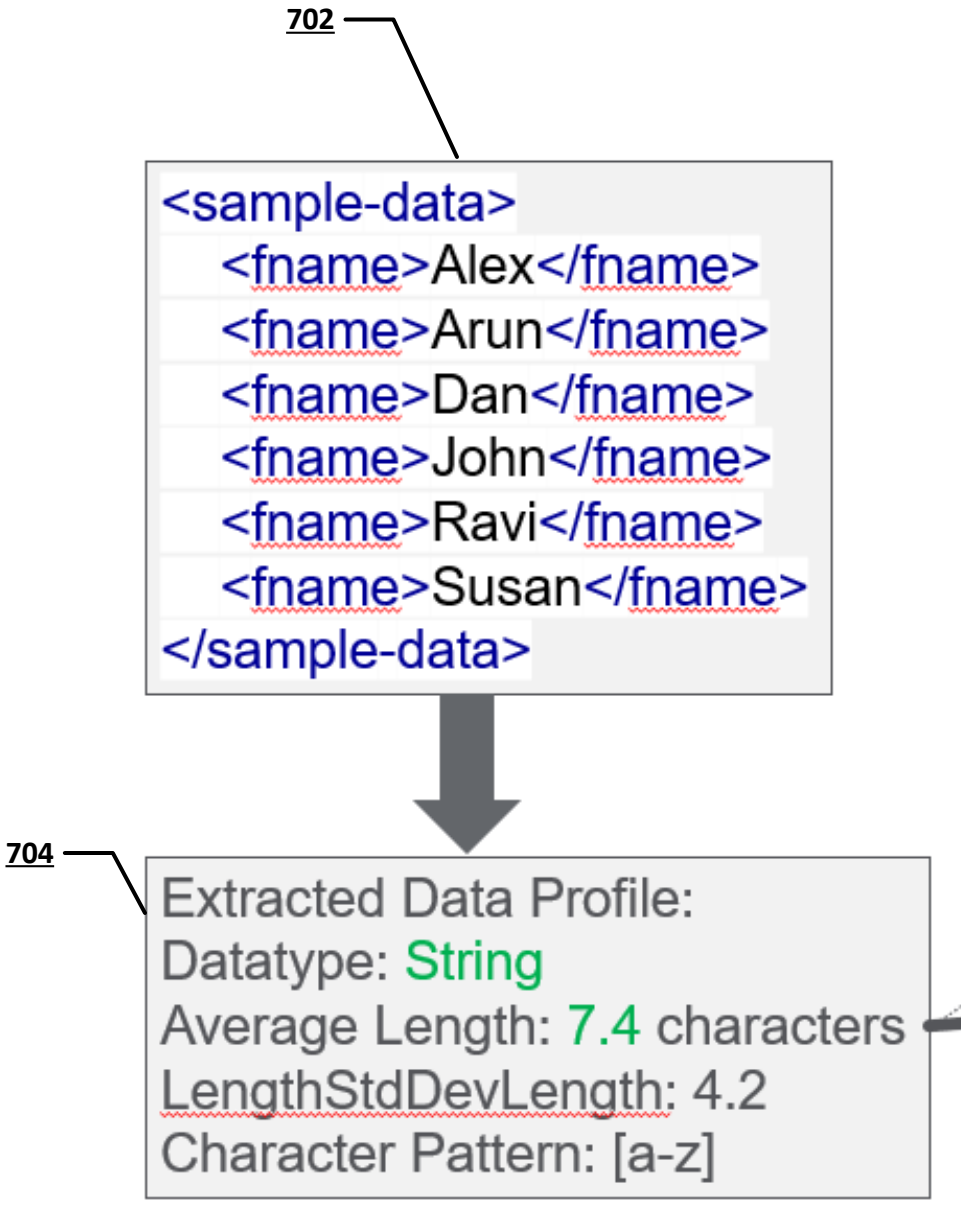

702
<sample-data>
<fname>Alex</fname>
<fname>Arun</fname>
<fname>Dan</fname>
<fname>John</fname>
<fname>Ravi</fname>
<fname>Susan</fname>
</sample-data>
704
Extracted Data Profile:
Datatype: String
Average Length: 7.4 characters
LengthStdDevLength: 4.2
Character Pattern: [a-z]

FIG. 7

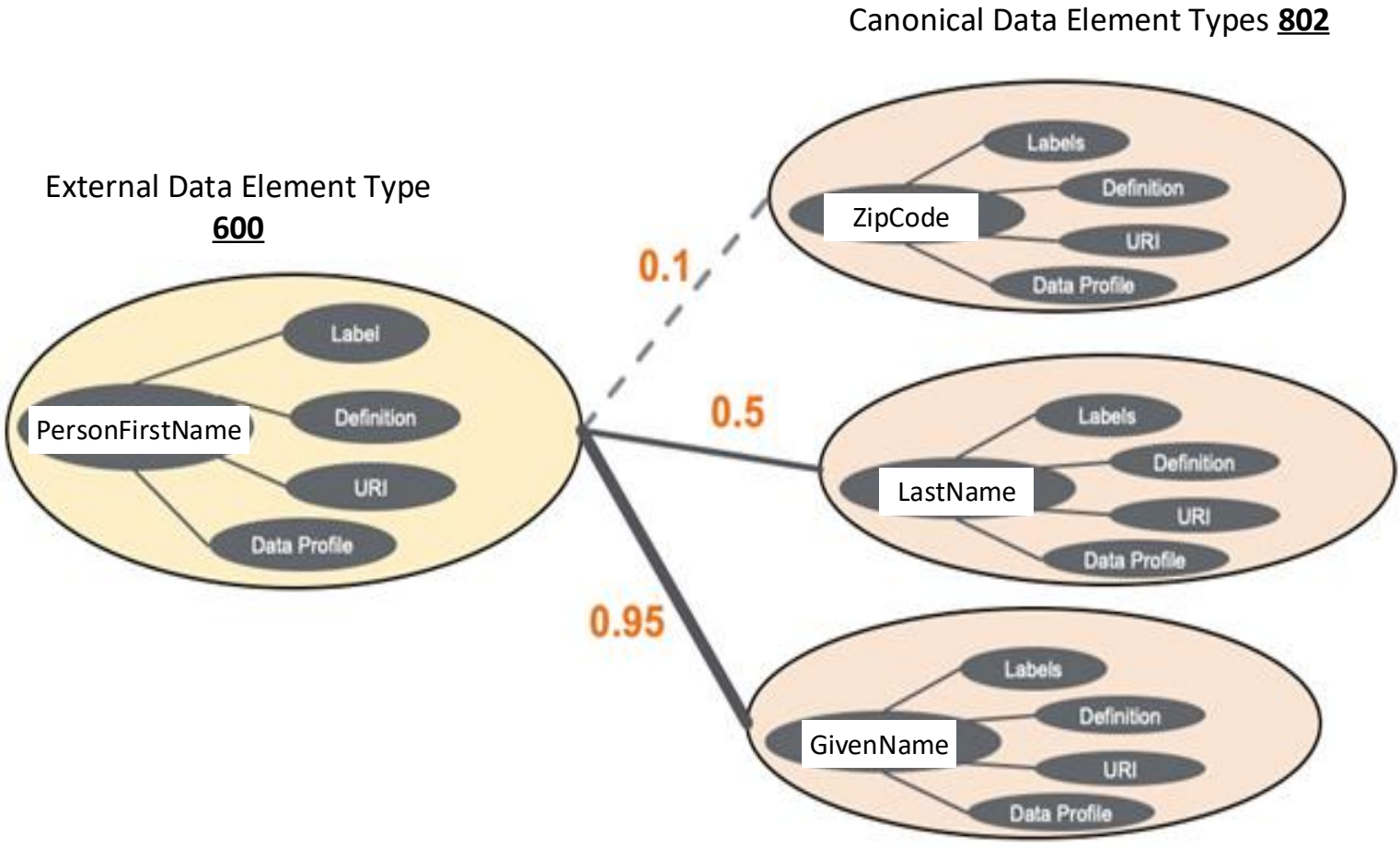
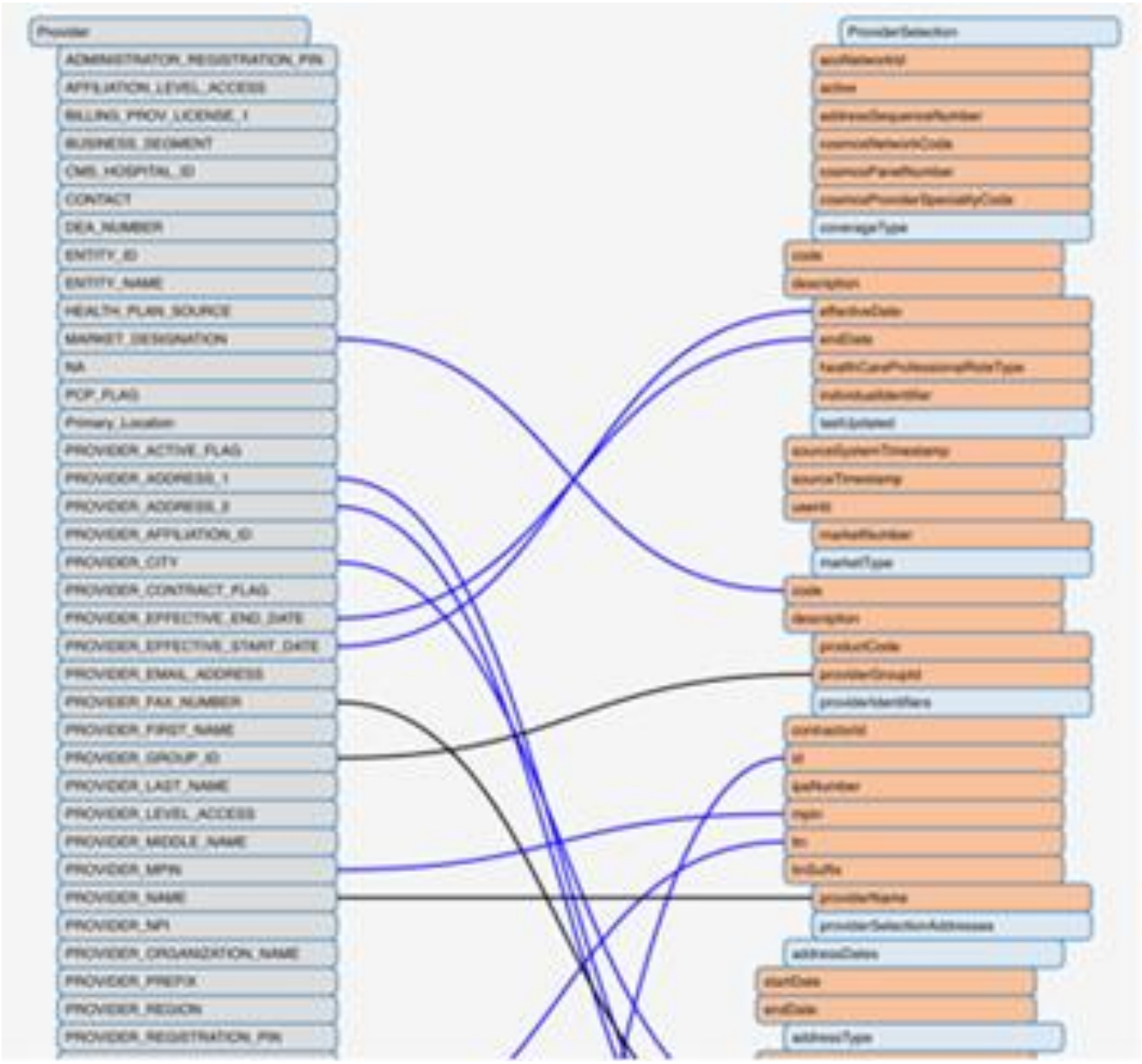

Canonical Data Element Types 802
External Data Element Type
600
Label
PersonFirstName
Definition
URI
Data Profile
0.1
0.5
0.95
Labels
ZipCode
Definition
URI
Data Profile
Labels
LastName
Definition
URI
Data Profile
Labels
GivenName
Definition
URI
Data Profile

FIG. 8

900



902

904

FIG. 9