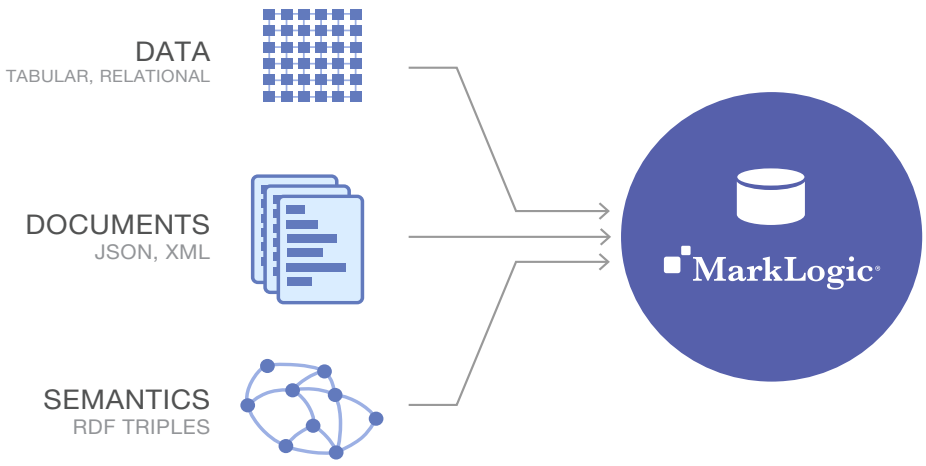
MarkLogic®

DATASHEET

Flexible Data Model

MarkLogic® is a multi-model database that provides native storage for JSON, XML, RDF, geospatial, and large binaries (e.g., PDFs, images, videos). With this approach, it is easy to get all of your data in and easy to make changes later on. Relational databases require predefined schemas and complex ETL to store data in rows and columns. With MarkLogic, you can load all of your data *as is*, without traditional ETL processes. Your structured and unstructured data—your data and metadata—can all be stored together in the same database. And, if you need to add another data source with a different schema later on, that is okay. In the words of one MarkLogic customer, MarkLogic's flexible data model "removes the shackles of relational technology."



The Document Model

At its core, MarkLogic stores and indexes data as JSON or XML documents. Among NoSQL databases, the document model is the most popular, and it helps solve many of the challenges with relational databases. Documents are ideal for handling varied and complex data, they are human-readable, they closely map to the conceptual or business model of the data, and they avoid the impedance mismatch problem that relational databases have.

Because the document model makes it possible to maintain and store multiple different schemas in the same database, data integration is easier and faster. And, with MarkLogic's powerful indexes, you can search and query across all of your data using your language of choice—whether it is JavaScript, XQuery, SQL, or SPARQL.

```
{
  "hospital": "Johns Hopkins",
  "operationType": "Heart Transplant",
  "surgeon": "Dorothy Oz",
  "operationNumber": 13,
  "drugsAdministered": [
    { "drugName": "Minicillan",
      "drugManufacturer": "Drugs R Us",
      "doseSize": 200, "doseUOM": "mg" },
    { "drugName": "Maxicillan",
      "drugManufacturer": "Canada4Less",
      "doseSize": 400, "doseUOM": "mg" },
    { "drugName": "Minicillan",
      "drugManufacturer": "Drugs USA",
      "doseSize": 150, "doseUOM": "mg" }
  ]
}
```

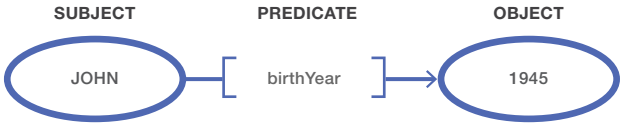
Example of a JSON document representing a surgical procedure at a hospital

MarkLogic®

Multi-Model With Semantics

Semantics describes MarkLogic's ability to store graph data as RDF Triples. Semantics enhances the document model by providing a smart way to *connect* and *enhance* the JSON and XML documents that MarkLogic stores, which is important for data integration and more powerful querying.

Semantics also provides *context* for your data. For example, consider a database that has information about parts, and one part is listed with a size of "42." But, where is the contextual information: *What are the units of "42"? What is the tolerance? Who measured it? When was it measured? Who can see this data?* That contextual information is the *semantics* of your data, and is easily stored in MarkLogic.




```
<http://dbpedia.org/resource/John>
<http://dbpedia.org/ontology/birthYear>
<http://dbpedia.org/resource/1945> .
```

Example of a semantic triple, which represents a fact about the world as a subject-predicate-object relationship. This example is from DBPedia, and is shown both graphically and expressed using the Turtle format.

Powerful and Composable

With MarkLogic's multi-model approach, you have the power to store and manage all of your data. And, it is fully composable—, a single query can run across any number of data types. The table below summarizes the usage and description of each data type in MarkLogic.

	JSON	XML	RDF	JSON/XML + RDF
Usage	Ideal for structured data stored as objects	Ideal for structured and unstructured data or text	Ideal for facts and relationships	Ideal for systems of data, text, and relationships
Description	• Schema-agnostic • Query with JavaScript • Compact and fast to parse • Six kinds of values: objects, arrays, floats, strings, booleans, nulls • Avoids namespaces, comments and attributes • Common data format for the web	• Schema-agnostic • Query with XQuery • Can store objects, sets, and many data types such as dates, durations, integers, and more • Uses namespaces (for embedding object types), comments, and attributes (for adding metadata) • More maturity than JSON as a data model	• Define entities and relationships • Atomic structure (cannot be broken down further) • Uses universal standards for data and querying (RDF and SPARQL) • Used for reference data, metadata, provenance	• Documents can contain triples • Triples can annotate documents • Graphs of triples can contain documents • Enhanced querying: - Expand a document search using graphs - Enhance graph search by linking to documents - Restrict document search using triples

About MarkLogic

MarkLogic is the world's best database for integrating data from silos, providing an operational and transactional Enterprise NoSQL database platform that integrates data better, faster, with less cost. Visit www.marklogic.com for more information.

MARKLOGIC CORPORATION
999 Skyway Road, Suite 200 San Carlos, CA 94070
+1 650 655 2300 | +1 877 992 8885 | www.marklogic.com | sales@marklogic.com