Problems in the use of big data

Structured data can be processed, but unstructured data cannot

There are many big data solutions, but no solution can process big data at the same level as humans can interpret them

We cannot use various kinds of data unrestrictedly in the field of business

Big data around us

Social media data

Information posted to social media by participants

Multimedia data


Audio and video data provided on distribution sites, etc. on the Web

Website data

Purchase histories, blog entries, etc. accumulated on EC sites, blogs, etc.

Customer data

Sales promotion data such as DM, membership card data, etc., managed in CRM systems

Big data

Sensor data


Travel records, temperature, acceleration, etc., which are detected by GPS, IC cards, RFID systems, etc.

Office data


Office documents and e-mails created on office PCs

Log data

Access logs, error logs, etc. automatically generated on WEB servers

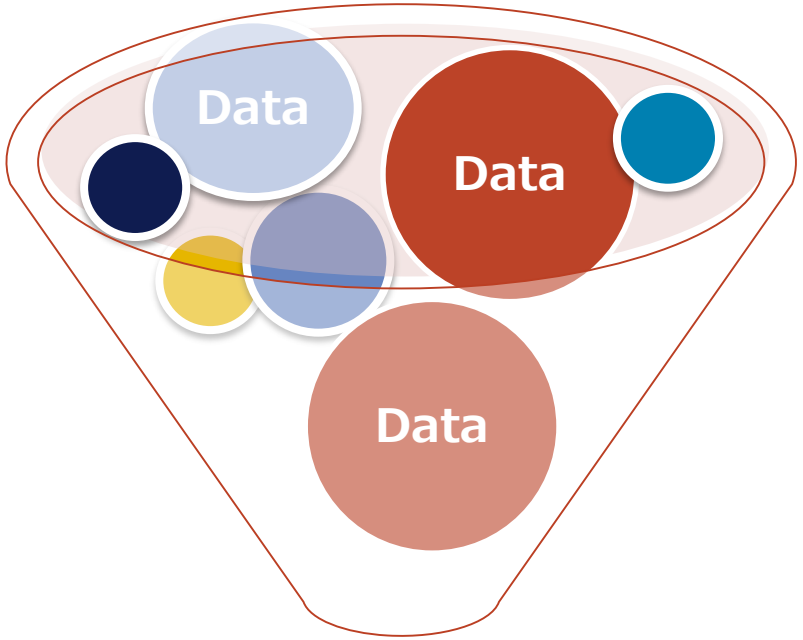
Operation data

POS data, transaction data, etc., generated on sales management operation systems

Source: Created by NTT Data based on materials published by the Ministry of Internal Affairs and Communications

NTT DATA

Variety – Utilize diversity –



Big data

Business side

For most businesses, it could be more significant to collectively process various kinds of data rather than just processing a large amount of data.

Difficult to realize due to high hurdles

Technology side

In handling various kinds of data, there still are restrictions on databases and peripheral technologies processing such data itself have not been well developed.

Ideal situation

Any data

As they are

Can take data in at any timing

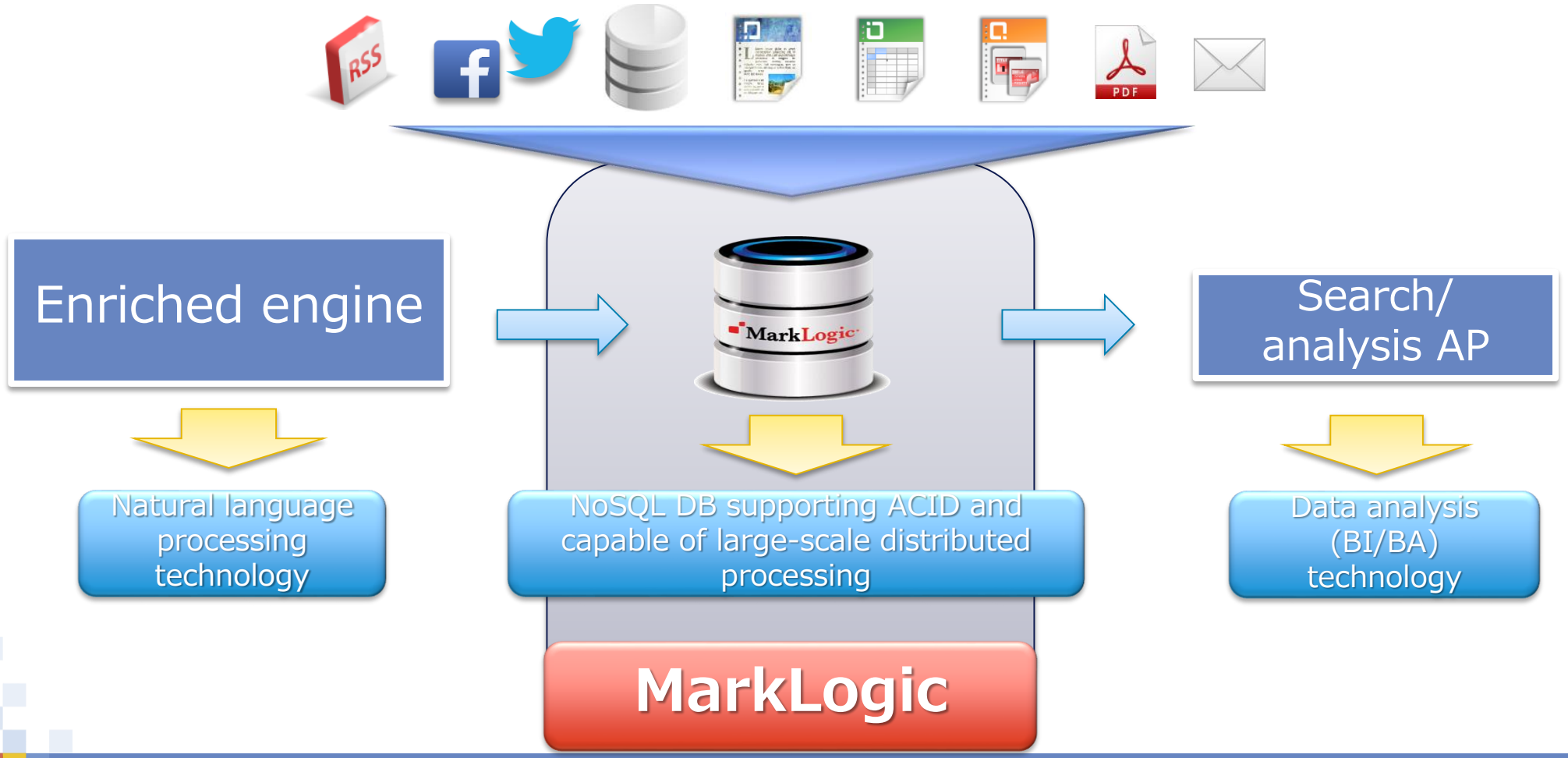
Use only data you want to use right now

- In the data structure you want
- Extract only necessary data
- Transform data to the format you want

⇒ Big data could also serve as a source to grasp changes in society and customers in the digital society.

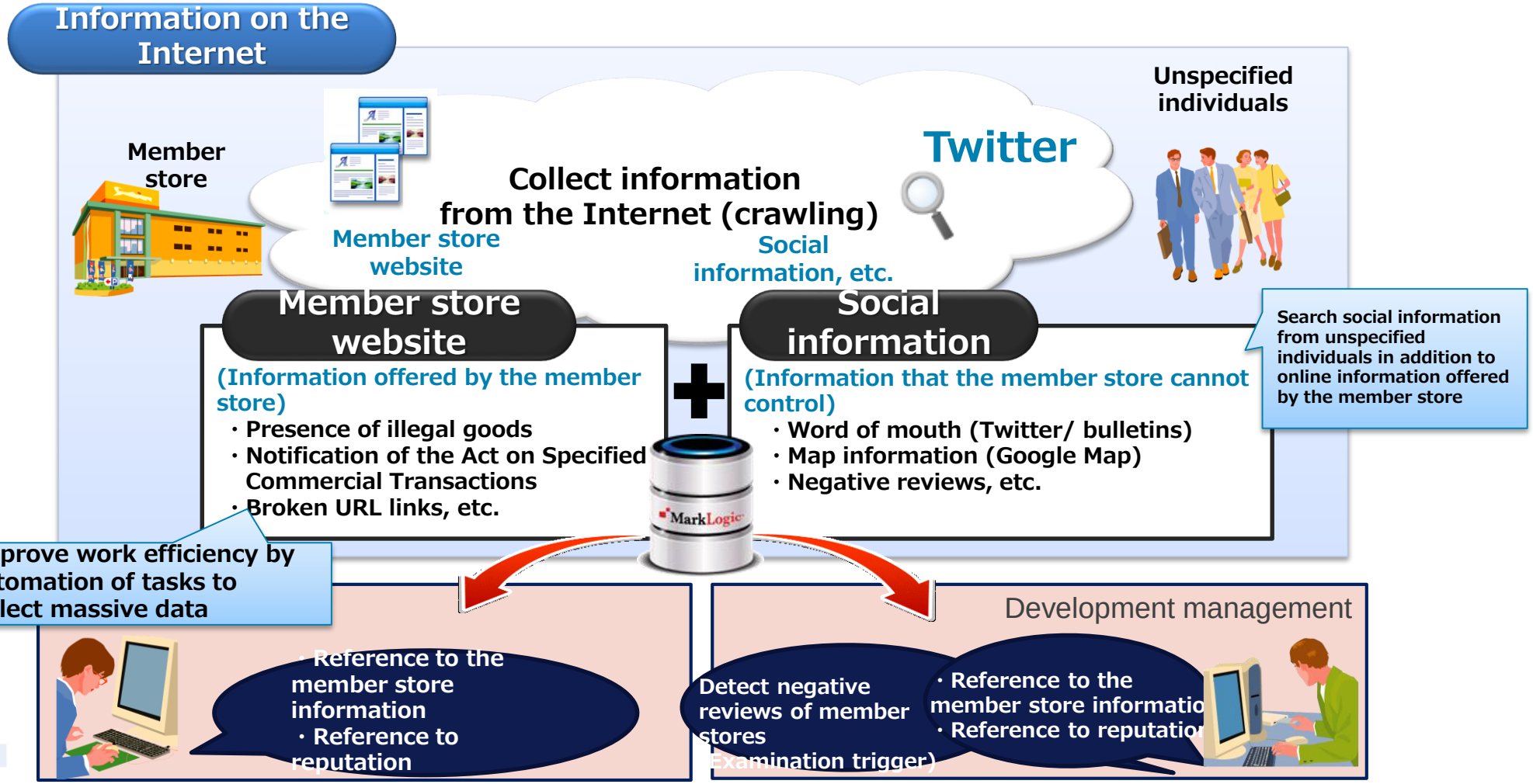
Use of MarkLogic Server as the infrastructure to realize services leveraging a wide variety of data

To build an unstructured data utilization infrastructure, we use MarkLogic Server as the core element of the highly flexible and expandable data management infrastructure which can manage data structures not initially expected, changes in data structures and unfixed-format data such as text documents in real time, search based on the meaning of data and complex relationships between data, in addition to fixed-format data managed by conventional RDB systems.



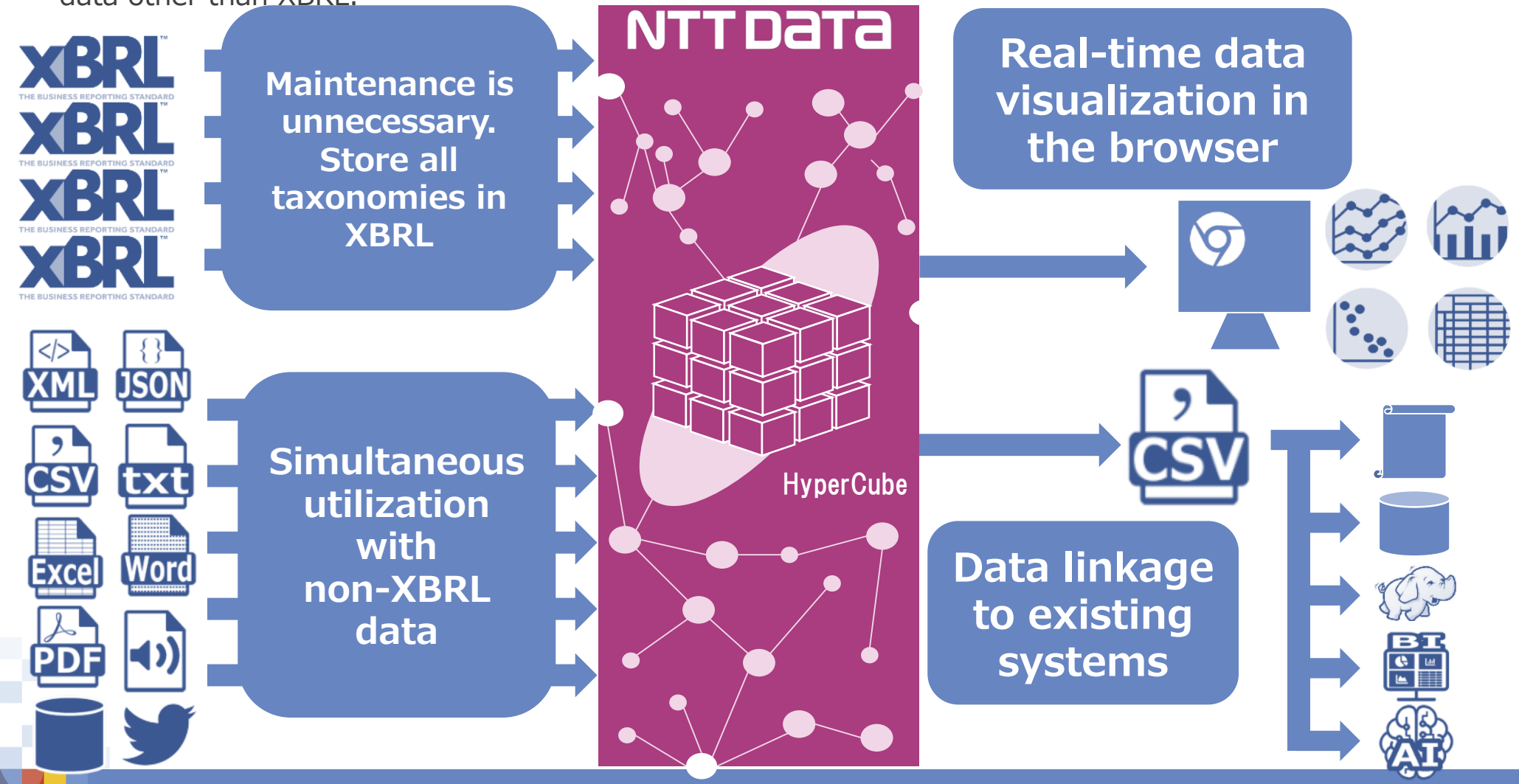
Concrete example of utilizing MarkLogic Server -Solution for member store examination-

This solution will increase the sophistication of the examination operation by customers, by conducting semantic interpretation and tagging for all information related to the examined store including information on EC sites and bulletin boards on the Internet. Then a variety of data will be stored in MarkLogic Server in a usable form.



Concrete example of utilizing MarkLogic Server -XBRL solution-

XBRL, which is used for exchanging international accounting data, is very complicated as it is. Also, XBRL files are hard to unrestrictedly process on RDBs as their formats are changed frequently. This solution enables users not only to process XBRL easily but also to collectively process a wide variety of data other than XBRL.



NTT DATA