MatDX: A Data Warehouse Solution for the Integration and Classification of Materials Information
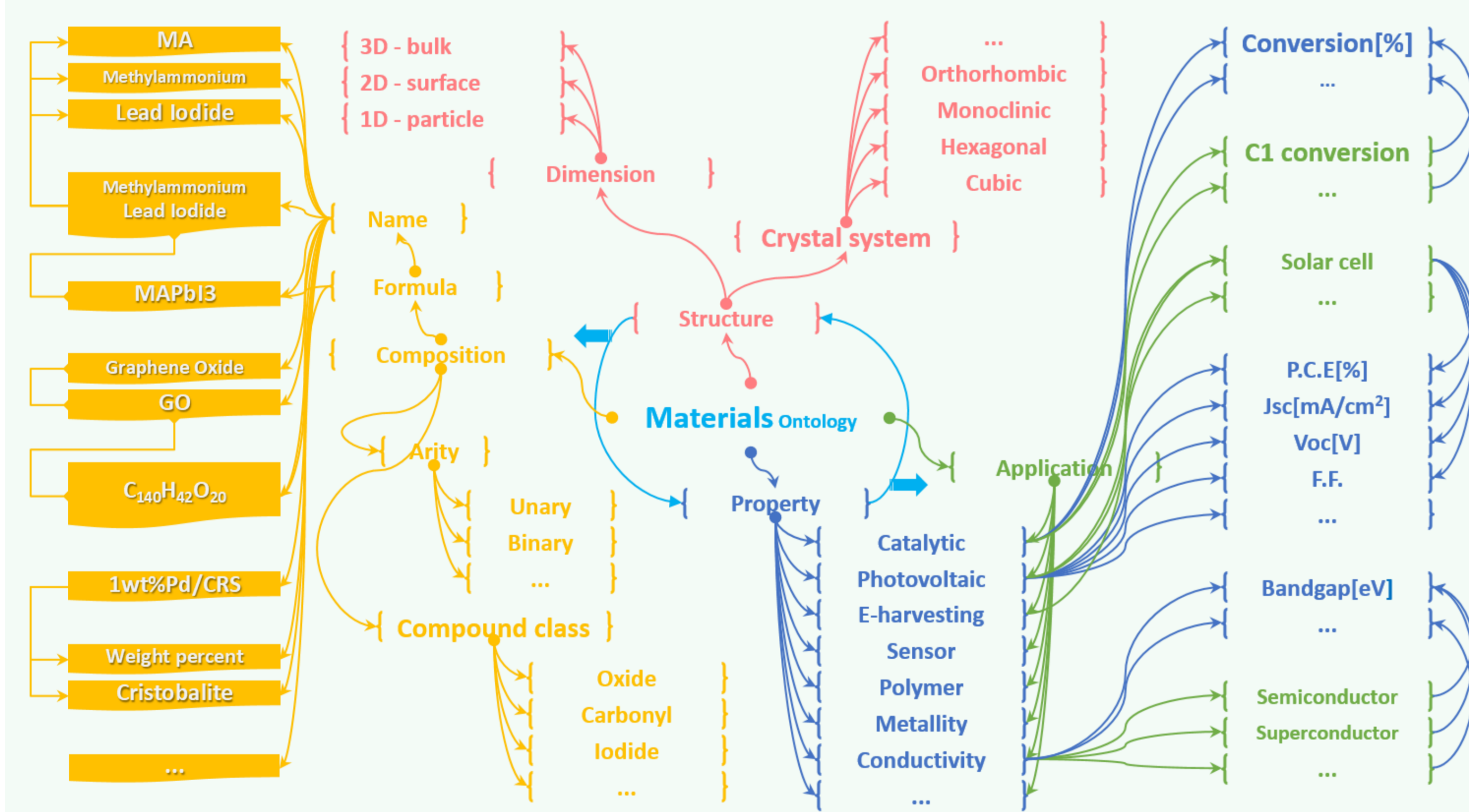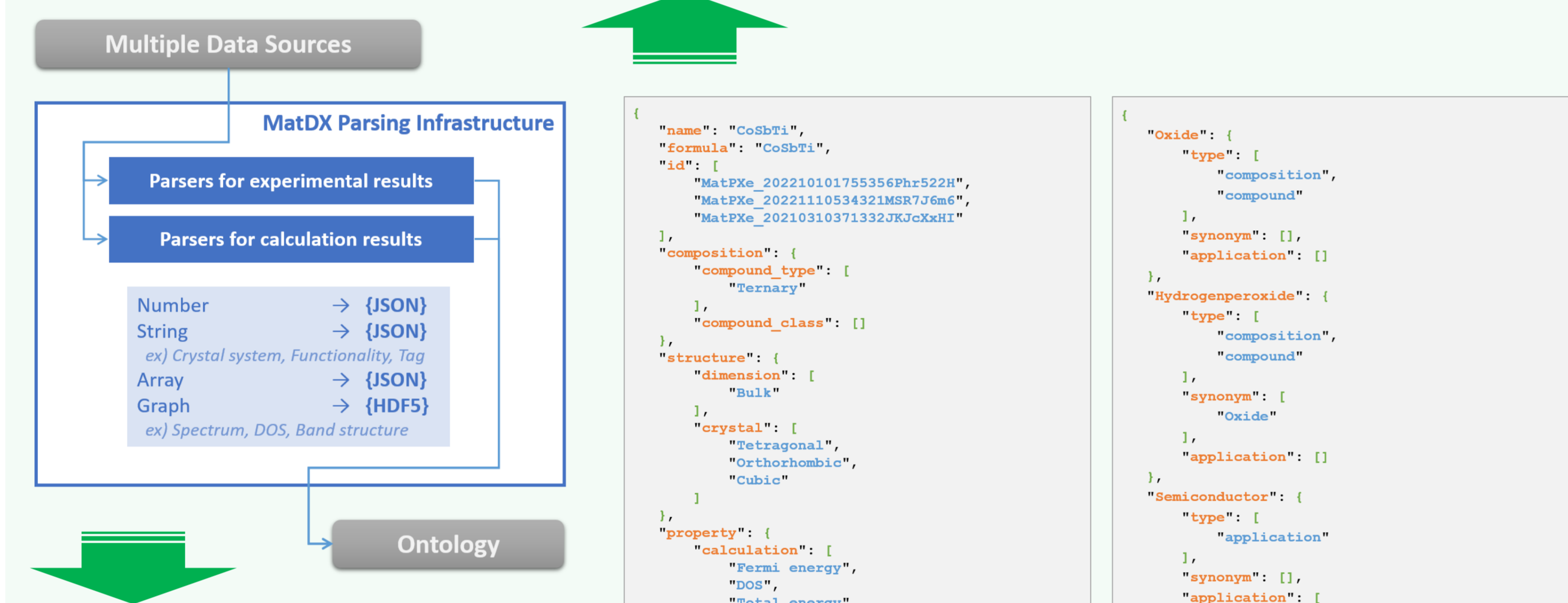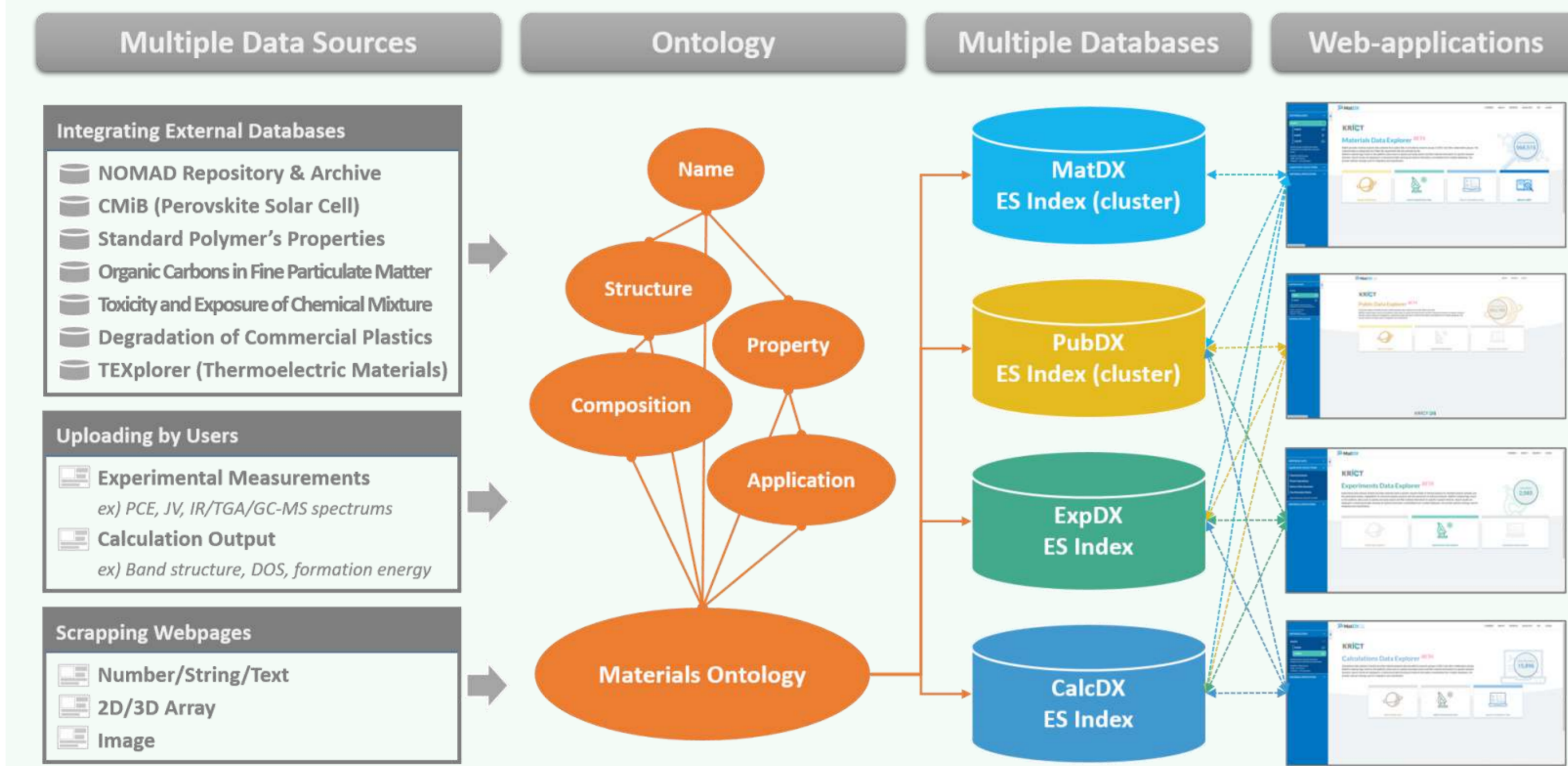
Jungho Shin[1], Hyunju Chang[1]

[1]Korea Research Institute of Chemical Technology, 141 Gajeongro, Yuseong, Daejeon 34114, Korea | jungho@krict.re.kr
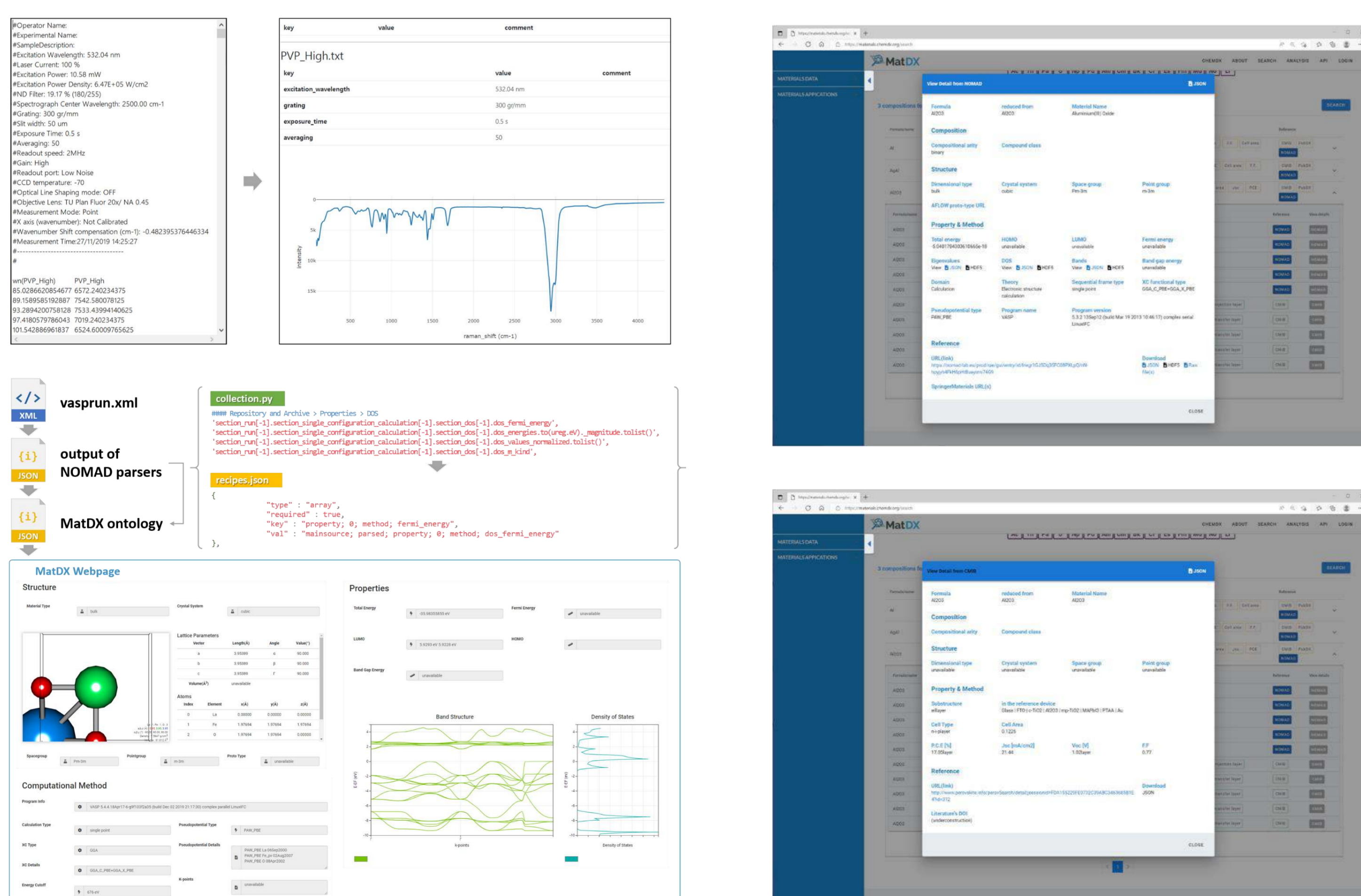
ABSTRACT In the field of data-driven materials science, the emergence of useful data infrastructures based on the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles, such as NOMAD, OPTIMADE, Materials Project, AFLOW, and OQMD, has accelerated this process even further. To facilitate the utilization of these data infrastructures, the integration and classification of numerous kinds of metadata related to materials properties are required. This, in turn, enables the research community to perform more accurate predictions through the use of machine learning algorithms. MatDX (Materials Data eXplorer) has been developed with a primary focus on the integration and classification of materials ontologies, including classes and instances in material name, composition, compound, structure, property, and applications. To facilitate integration, MatDX has been developed as a kind of data warehouse solution, which serves as a useful strategy for connecting multiple databases. Material tags have been introduced through classification by using the materials ontologies, and they are used to search and view detailed information for interesting materials in an easy and quick way. Further analysis functionality to visualize statistically meaningful relationships has been served in an interactive way.

The ultimate goal of these services is to enable researchers to discover novel materials for target properties based on numerous research data. MatDX, including three sub-categories of PubDX (for published data), ExpDX (for experimental data), and CalcDX (for calculated data), is ready to be accessed through the address of http://materials.chemdx.org.
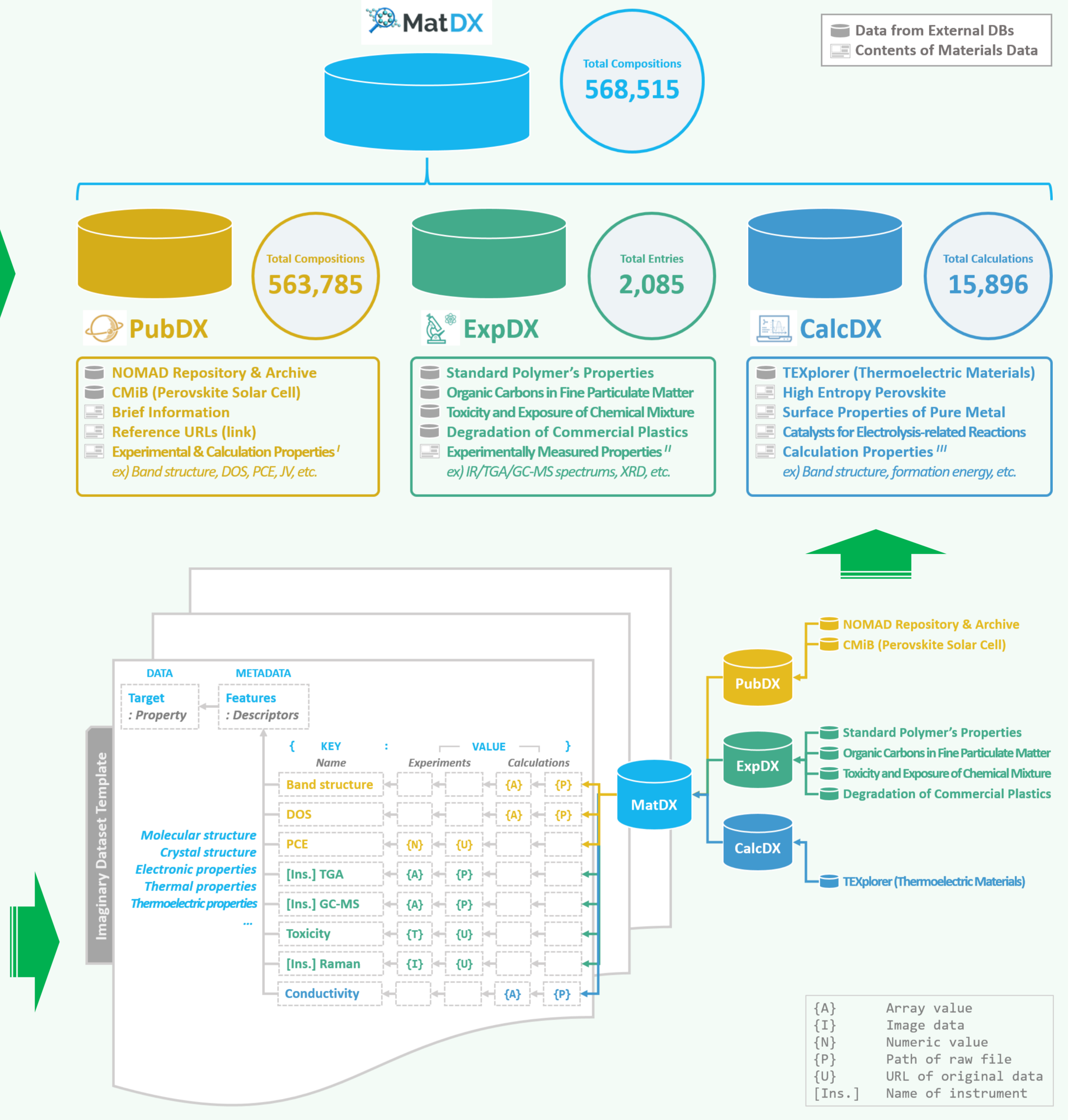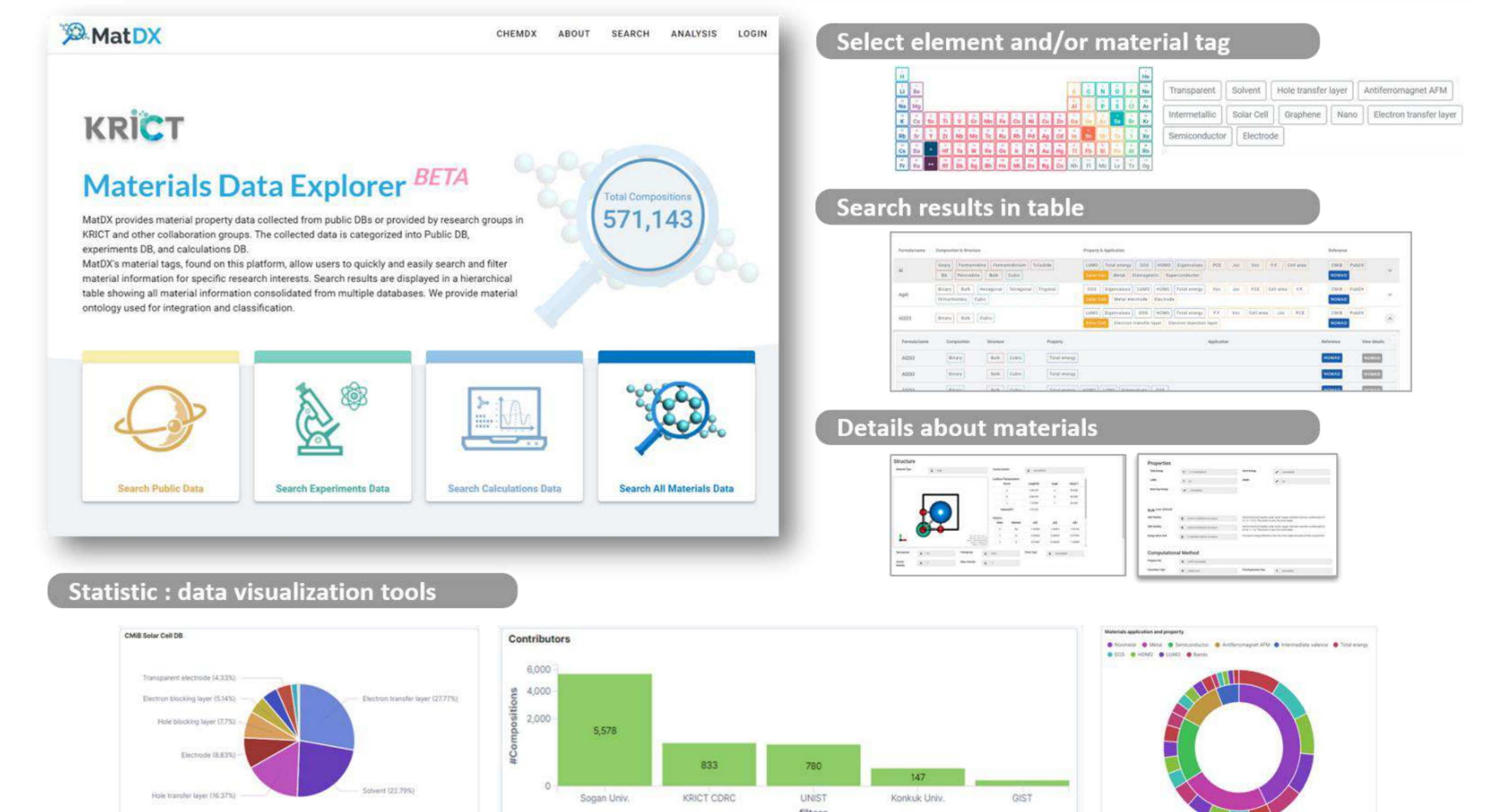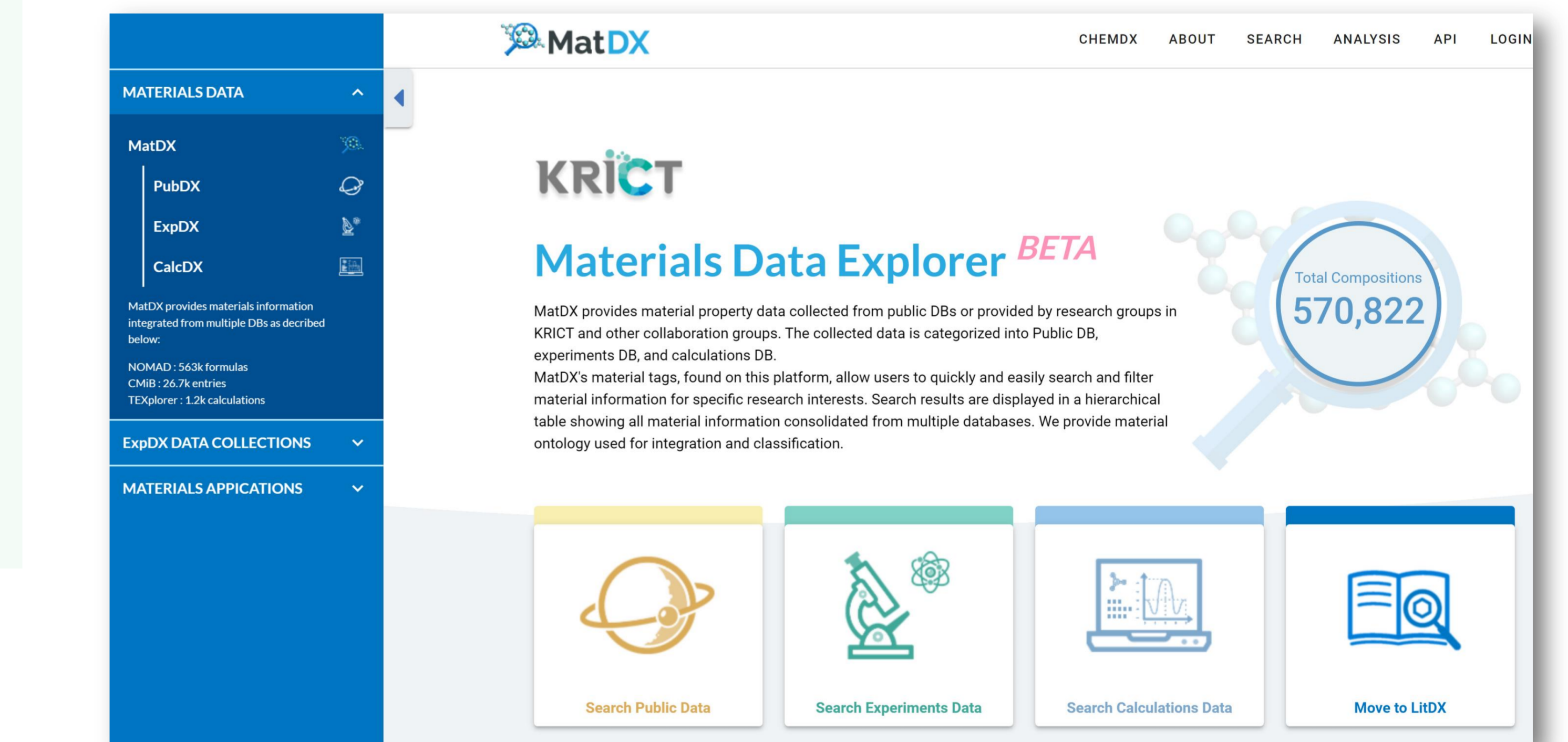
DATA PROCESSING

STRUCTURE

ONTOLOGY

WEB SERVICE - HOME

DATA VISUALIZATION

FUNCTIONALITIES