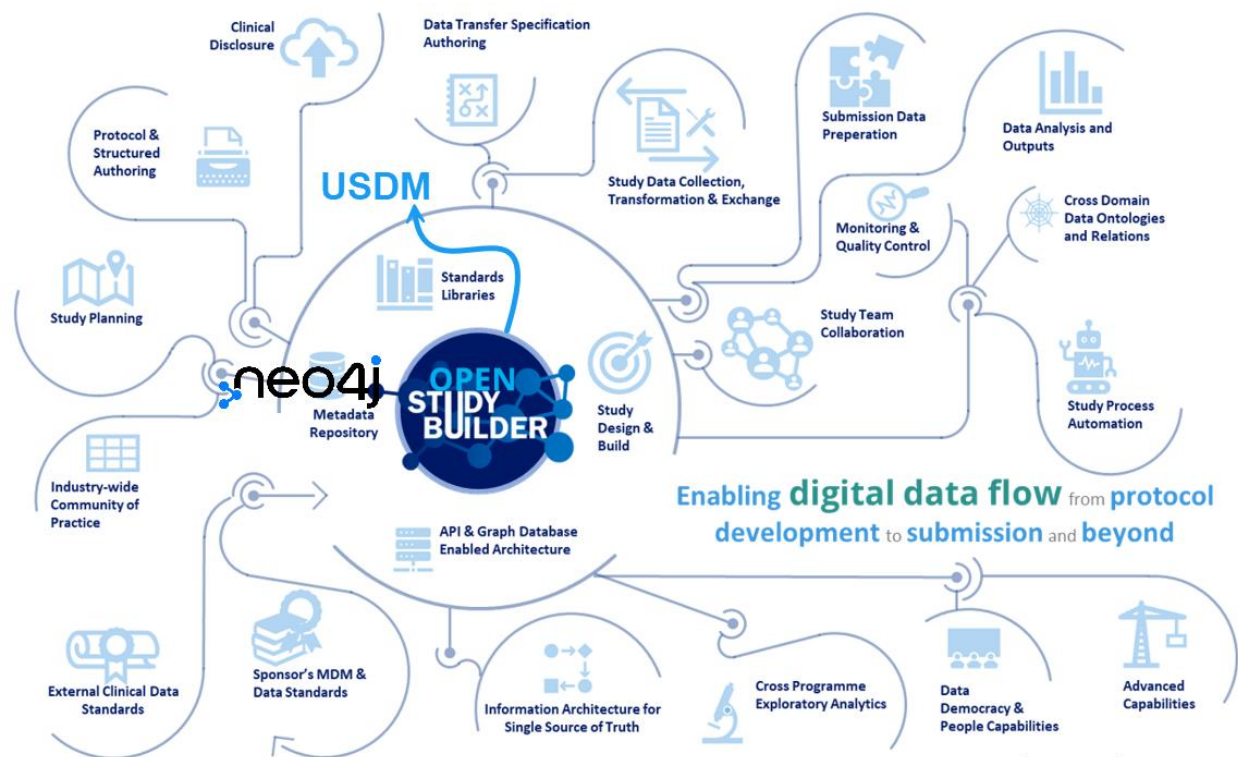


Leveraging the USDM Standard in OpenStudyBuilder for Enhanced Data Exchange

Marius Conjeaud, Neo4j, Bordeaux, France
Nicolas de Saint Jorre, Novo Nordisk, Nantes, France

ABSTRACT

The OpenStudyBuilder (OSB) Metadata Repository (MDR) and Study Metadata Repository (SDR) serve as pivotal components in Novo Nordisk's comprehensive approach to digitizing the entire clinical process. This encompassing digital transformation commences with the Protocol and extends to the Trial Files and Listings (TFL), facilitated by the Schedule of Activity. The OSB leverages these repositories to drive the production of the Case Report Form and specify the Tabulation of Metadata. Central to this process is the Activity concept, which orchestrates the interconnection of metadata elements through a Graph database provided by Neo4j. A recent milestone achieved by the OSB is the capability to extract Protocol metadata within the Unified Study Definition Model (USDM), enabling the generation of a Novo Nordisk protocol but at the same time, an ICH - M11 one, by leveraging shared metadata.



INTRODUCTION

In the realm of clinical research, the digitization of processes has become paramount in enhancing efficiency, accuracy, and overall management of trials. For Novo Nordisk, this digital transformation begins at the very foundation of clinical studies, the Protocol, and extends throughout the entire lifecycle, encompassing Trial Files and Listings (TFL). The Schedule of Activity is driving the specification of the Case Report Form but also of the tabulation to SDTM for example. This interconnected ecosystem is further empowered by the Activity concept, which intricately binds various metadata elements with the aid of a Graph database. Notably, a recent advancement in this framework allows the OpenStudyBuilder to extract Protocol metadata within the Unified Study Definition Model (USDM), thereby facilitating the generation of protocols compliant with the ICH - M11 standard, as Sponsor version through the utilization of shared metadata. This introduction sets the stage for understanding the critical role played by the OSB MDR and SDR in driving Novo Nordisk's comprehensive approach to digitizing the clinical process from end to end.

WHAT IS THE OPENSTUDYBUILDER ?

The OpenStudyBuilder (OSB) is an open-source project for clinical study specifications. This tool is a new approach for working with studies that once fully implemented will drive end-to-end consistency and more efficient processes - all the way from protocol development and CRF design - to creation of datasets, analysis, reporting, submission to health authorities and public disclosure of study information.

The heart is the Metadata Repository using a graph database containing a biomedical concept. This allows for FAIR data, semantics, traceability and automation. Then there is the Application itself where all data can be browsed, entered, updated and managed. And finally, there is a very powerful API which provides the complete functionality through a standard interface!

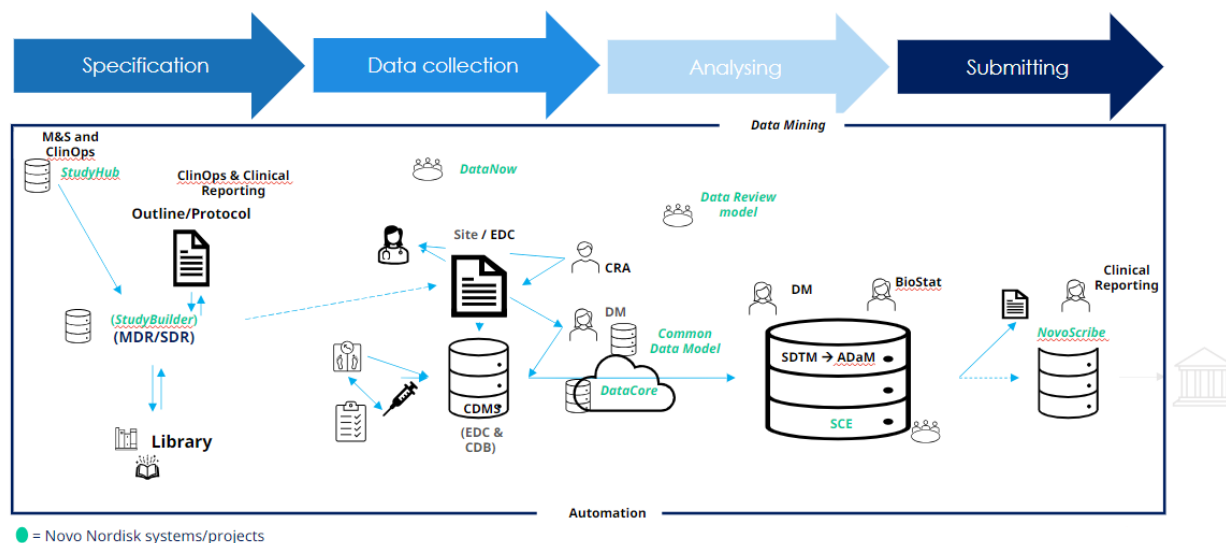


Figure 1: Global integration of the OpenStudyBuilder / Neo4j database in the ecosystem at Novo Nordisk

THE LIBRARY

In the library section you can browse all kinds of standards, like for CDISC the SDTM, ADAM, SEND, CDASH and many more. Also, other dictionaries like SNOMED, LOINC, UCUM is available - MedDRA could be loaded as well of course. Want to see differences of different SDTM versions? No Problem with the OpenStudyBuilder! As everything is FAIR, concepts can be created and maintained, so these can easily be reused. A core feature is the usage of syntax templates, for example for objectives and endpoints. Here general standard wordings and types are defined, which can then be instantiated with a concrete compound or comparator.

THE STUDY AREA

All study-related content which is needed for the protocol and Trial Domains can be managed in the study area. Objectives, endpoints, study design, population - whatever is needed can be entered. There is also a nice protocol flowchart. Lately, we were adding here, a Unified Study Definition Model (USDM) API converting our metadata into a full JSON file following the specification.

THE API

The API provides a standard interface which is very powerful. The application itself uses the API to update the database. Everything could also be done by using the API in a different fashion. Notable is also the USDM API Adapter which supports the TransCelerate Digital Data Flow (DDF) connections.

THE DATABASE

The data model to support this tool is very densely connected and must have a high level of flexibility because it is still growing and will keep growing for years as we extend it. It also includes versioning and auditing of all manipulated metadata, which increases the complexity and connectedness of the model.

This is why we chose a graph database (Neo4j) over other database types, because it is the only technology that can cater for such levels of connectedness.

WHAT IS THE DDF – USDM ?

The TransCelerate Digital Data Flow project (DDF) aims to a future state of fully automated, dynamic, study start-up readiness. This automated end-to-end process is also the goal of the OpenStudyBuilder. In the context of the DDF, the OpenStudyBuilder can be seen as an SDR (Study Definition Repository), but with more capabilities than what DDF is containing in its first version.

The main advantage of DDF is the standard API which allows the connection of very different up- and downstream systems. All tools which support the DDF project through their standard APIs should also be usable with the OpenStudyBuilder, which supports the same APIs.

THE API – USDM ENDPOINT

The OSB provide a dedicated API endpoint that is only requesting the reference of the Study and will return a complete USDM JSON file based on the version 3 of the DDF definition (see figure 2).

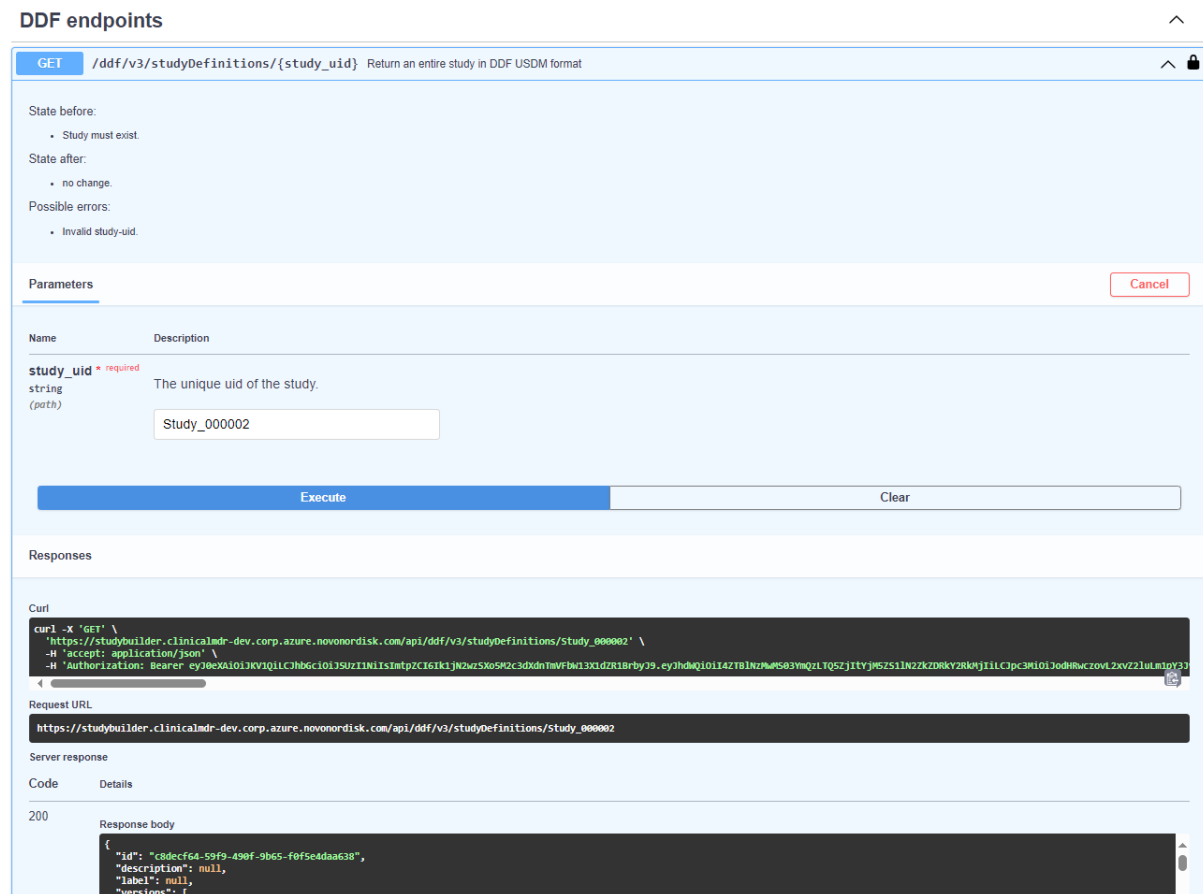


Figure 2: DDF api endpoint showing the USDM JSON extraction for a dedicated study

THE USDM IN THE FRONT APPLICATION

With the utilization of USDM version 3, the front end of the OpenStudyBuilder now boasts the capability to display the JSON view within the Study level section, allowing for direct file download through a dedicated button. Once the minimum set of metadata is established for a specific study, users can seamlessly access the USDM viewer by navigating to the View Specification menu, as illustrated in Figure 3.

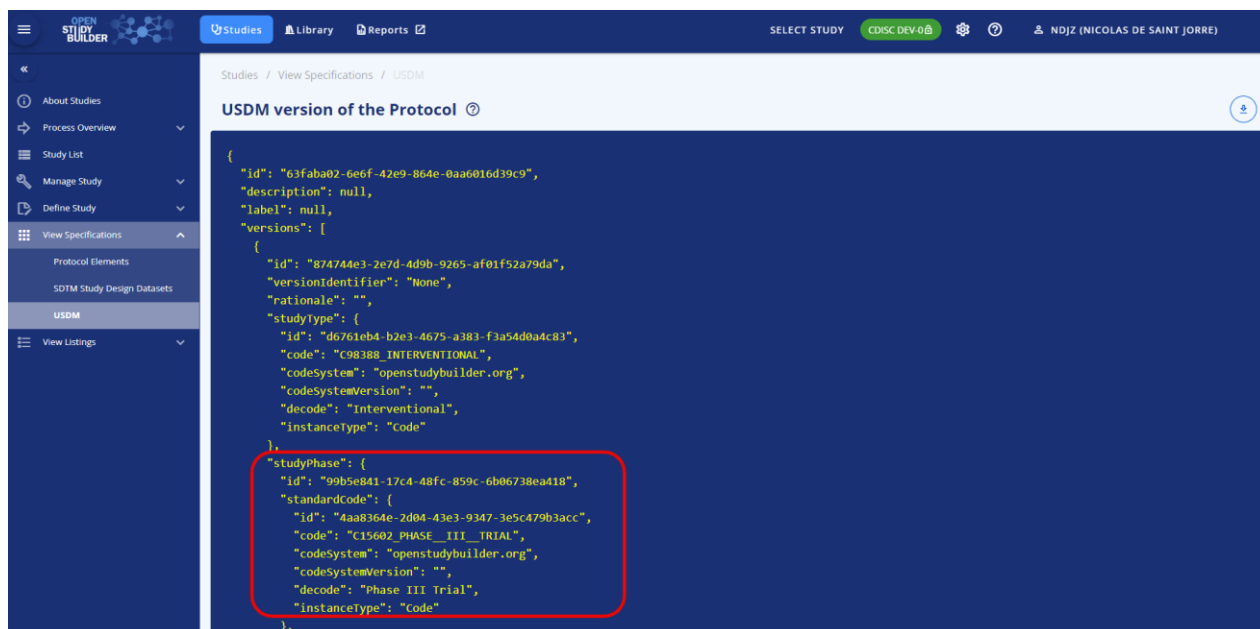


Figure 3: USD view inside the OpenStudyBuilder application

EXPORT FEATURE BUT LOOKING FOR IMPORTING

The Open Source team looking to enhancing the capabilities of the OpenStudyBuilder is willing to integrate a USD importer endpoint. This strategic initiative aims to streamline and fortify the platform's functionality by enabling seamless integration of data aligned with the Unified Study Definition Model (USD). With the addition of this importer endpoint, users will gain the ability to effortlessly incorporate USD-compliant data into the OpenStudyBuilder ecosystem, further solidifying its position as a versatile and adaptable solution for managing clinical research processes.

INTEGRATION OF A M11 TEMPLATE

The incorporation of an ICH - M11 template within the OpenStudyBuilder would represent a significant advancement, as it leverages the USD JSON metadata to seamlessly generate an HTML version of the M11 protocol that will be converted into a PDF doc. By harnessing the power of USD-aligned data, this functionality enables the production of M11 documents that align with industry standards, providing a parallel reference alongside the Sponsor version. This integration not only enhances efficiency and accuracy but also ensures compliance with regulatory requirements, thereby empowering researchers and stakeholders with a comprehensive and standardized view of the protocol within the clinical research landscape.

By harnessing the power of USD-aligned data, this functionality not only streamlines the production of M11 documents but also ensures that they adhere to industry standards. This capability provides a parallel reference alongside the Sponsor version, enhancing efficiency, accuracy, and compliance with regulatory requirements. Ultimately, this integration empowers researchers and stakeholders by offering a comprehensive and standardized view of the protocol within the clinical research landscape, thereby optimizing the document management process. And the idea is here to produce the Sponsor protocol as a Word Document as a Comment Protocol Template (CPT from TransCelerate) as the USD format and then the M11 one. See figure 4 and 5 below.

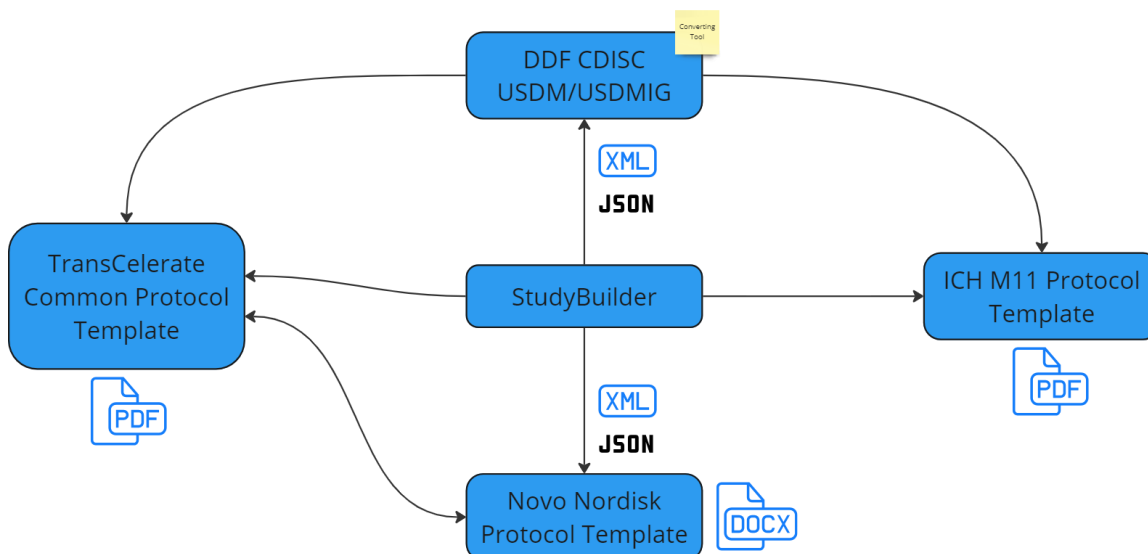


Figure 4: How is the OpenStudyBuilder going to manage multiple targets with one set of metadata

ICH M11 Template

Coming from the OpenStudyBuilder

Protocol Full Title:	[Protocol Full Title] The protocol should have a descriptive title that identifies the scientific aspects of the trial sufficiently to ensure it is immediately evident what the trial is investigating and on whom, and to allow retrieval from literature or internet searches.
Sponsor Confidentiality Statement:	[Sponsor Confidentiality Statement] Insert the Sponsor's confidentiality statement, if applicable, otherwise delete.
Protocol Number:	[Protocol Number] A unique alphanumeric identifier for the trial, designated by the Sponsor, is a standard part of trial data, and should be included for most trials.
Version:	[Version] An optional field for use by the Sponsor at their discretion.
Amendment Number:	[Amendment Number] Enter the amendment number. If this is the original instance of the protocol, indicate Not Applicable.
Amendment Scope:	[Amendment Scope] [Country/Region Identifier] Acceptable entries for amendment scope are: "global" or "Country-specific/Regional" Use the ISO-3166 region or country identifier (for example, DE or EU). For global trials delete the Country/Region Identifier field.
Compound Number(s):	[Compound Number] Enter the Sponsor's unique identifier for investigational compound(s) in the trial. Add or delete additional fields as needed.

Figure 5: ICH – M11 Template to be integrated into the OpenStudyBuilder

REGULATORY INTEGRATION

The integration of the ICH - M11 template into the OpenStudyBuilder represents a significant advancement in the realm of protocol generation. This integration not only facilitates the creation of protocols but also establishes a smooth pathway for their dissemination to regulatory agencies in a manner that is both direct and standardized. Through harnessing USDM-aligned data to generate compliant HTML versions of the M11 protocol, this feature guarantees that the sharing of protocols with regulatory bodies becomes not only straightforward but also universally applicable. As a result, this functionality serves to enhance transparency, compliance, and efficiency in all interactions with regulatory authorities.

Furthermore, the Food and Drug Administration (FDA) has articulated its expectation to receive protocols in the M11 format. This underscores the critical nature of this integration, as it not only aligns with regulatory expectations but also ensures seamless compliance with the FDA's requirements, thereby streamlining the submission process and bolstering the overall efficiency of regulatory interactions.

A GRAPH DATABASE TO RULE THE BIOMEDICAL CONCEPT

A graph database such as Neo4j provides an ideal infrastructure for managing the metadata of a Metadata Repository (MDR) and facilitating the connection of biomedical concepts across various nodes. By leveraging the graph model, Neo4j enables the representation of complex relationships between different types of metadata as interconnected nodes and edges. In the context of biomedical concepts, this means that each concept can be represented as a node, and relationships between concepts can be defined as edges, allowing for the seamless connection and traversal of these interconnected concepts. This approach not only supports the storage and retrieval of metadata but also enables the exploration of relationships between different biomedical concepts, thus offering a powerful means to analyze and understand complex biomedical data in a highly interconnected manner. The graph database's ability to handle such intricate relationships and metadata connections makes it an invaluable tool for supporting biomedical research, knowledge discovery, and data-driven decision-making in the healthcare domain.

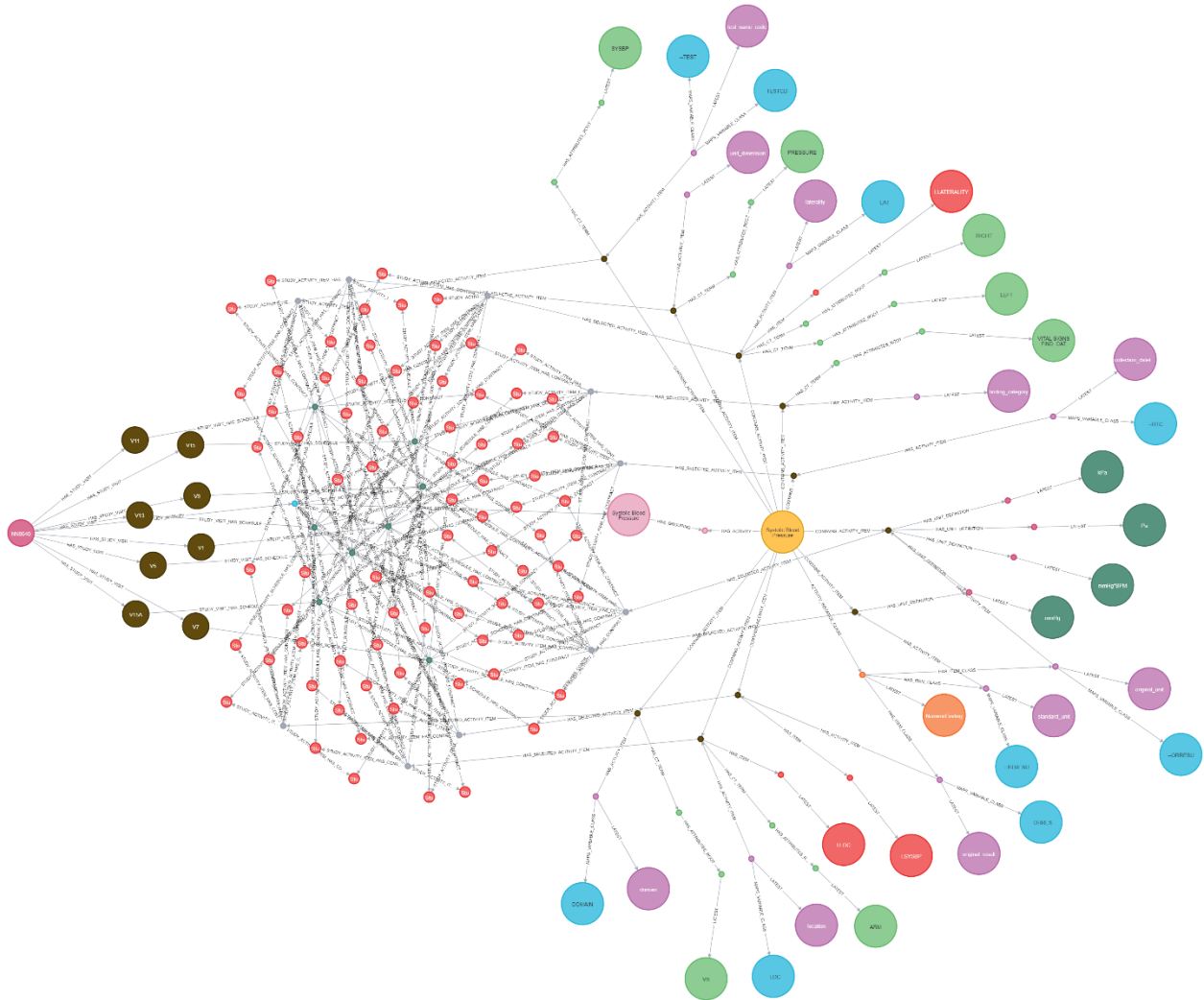


Figure 6: Example of the Systolic Blood Pressure ActivityInstance inside the Neo4j graph database

CONCLUSION

The successful generation of JSON aligned with the Unified Study Definition Model (USDM) underscores the efficacy of the OpenStudyBuilder's approach in managing complex metadata. This achievement underscores the pivotal role of the Neo4j Graph database in orchestrating the intricate interconnections within our metadata ecosystem. As we look to the future, our focus will center on integrating an M11 display to further strengthen our capacity to produce protocols compliant with industry standards. Furthermore, the planned incorporation of a USDM import function represents a pivotal next step, promising to enhance our ability to seamlessly integrate and leverage USDM-aligned data within our ecosystem.



Figure 7: View of the Activity Concept Model inside the graph database – Heart of the USDM conversion

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name: Marius Conjeaud

Company: Neo4j

Email: marius.conjeaud@neo4j.com

Co-author Name: Nicolas de Saint Jorre

Company: Novo Nordisk

Email: ndjz@novonordisk.com

Website: openstudybuilder.com

Brand and product names are trademarks of their respective companies.