

**Project title:** Implications of Medical Low Dose Radiation Exposure

**Grant Agreement:** 755523

**Call identifier:** NFRP-2016-2017

**Topic:** NFRP-9

**Documentation of the Sparklis tool**

**Lead partner:** Inserm

**Author(s):** Bernard Gibaud

**Work Package:** WP2

**Estimated delivery:** 28 September 2020

**Actual delivery:** 28 September 2020

**Type:** Report

**Dissemination level:** Public

**Version :** V1.0

1. Introduction

This document provides a basic documentation of the Sparklis tool, a tool allowing MEDIRAD end-users to navigate in the Semantic database of the Image and Radiation Dose BioBank (IRDBB) and to build SPARQL queries. This semantic database is represented in an RDF graph.

For an introduction to the IRDBB semantic ddatabase, see the presentation available at: https://eibir.teamwork.com/#/files/4620923

History of versions

|  |  |  |
| --- | --- | --- |
| **Version** | **Date** | **Description** |
| V1.0 | 28/09/2020 | First version corresponding to the distribution of the Semantic translator package SEMANTIC\_TRANSLATOR\_TAG = 0.0.70 containing the SEMANTICTRANSLATOR\_VERSION="0.8.10" based on ONTOLOGY\_VERSION="1.3.15" |

1. Principle of the Sparklis tool

### Why Sparklis ?

*SPARQL Protocol and RDF Query Language* (in short *SPARQL*) is the language defined by the W3C to query RDF graphs. SPARQL 1.1 is both a language for expressing queries and a protocol to communicate those queries and the answers to those queries [1].

As other query languages (e.g. SQL for querying relational databases), SPARQL is quite complex, and can hardly be used by end-users, who generally neither know the details of the syntax and the semantics of the language, nor are aware of the detailed structure of the RDF graph to be queried. This is why the designers of IRDBB specified a number a predefined queries, that can be directly selected from the IRDBB\_UI web server.

However, there are many situations where end-users will need to design their own queries.

Sparklis allows addressing this need by assisting users in building their own queries, interactively, without having to know the syntax of the language nor the detailed structure of the RDF graph.

### Principle

The principle of Sparklis is to enable the user navigating in available RDF data. For achieving this goal the Sparklis tool transparently explores the graph by accessing to it through a SPARQL endpoint, and provides the user information about available data and relations characterizing a particular class. From this information the user can select what information should be specified in the query so that it can be retrieved. The overall (SPARQL) query is built progressively, and translated to the user in natural language. More details about the functioning of Sparklis are provided in [2].

Sparklis was designed by Sébastien Ferré from IRISA in Rennes. Sparklis was adopted by CNRS for its PERSEE bibliographic database [3].

1. Use of the Sparklis tool

The Sparklis tool can be selected from the IRDBB\_UI interface by clicking on the Sparklis button (Figure 1), which triggers the opening of the Sparklis screen (Figure 2). This screen has three main parts:

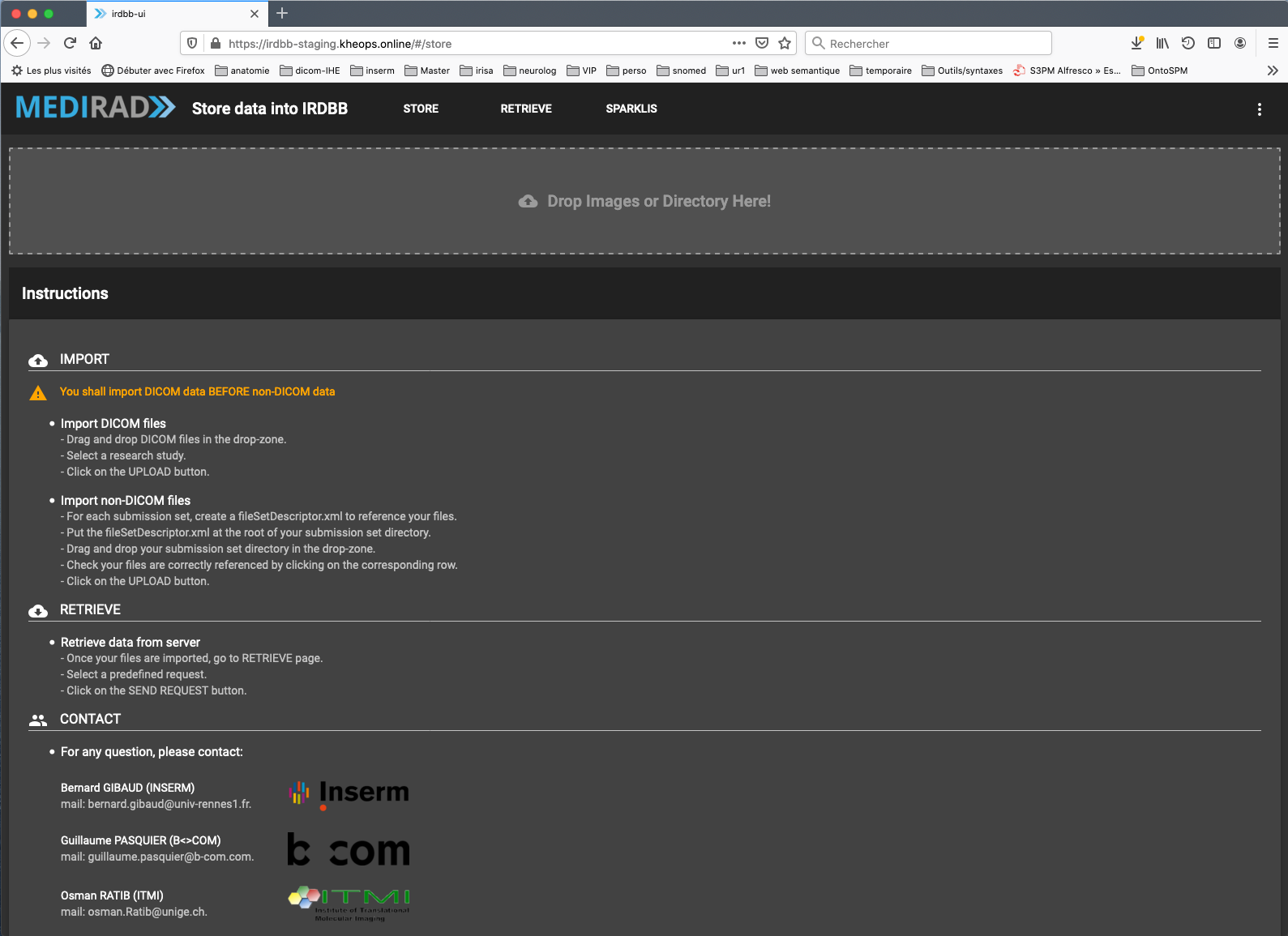


Figure 1: Main screen of IRDBB\_UI (web server)

(Sparklis button in the upper part, in the middle)

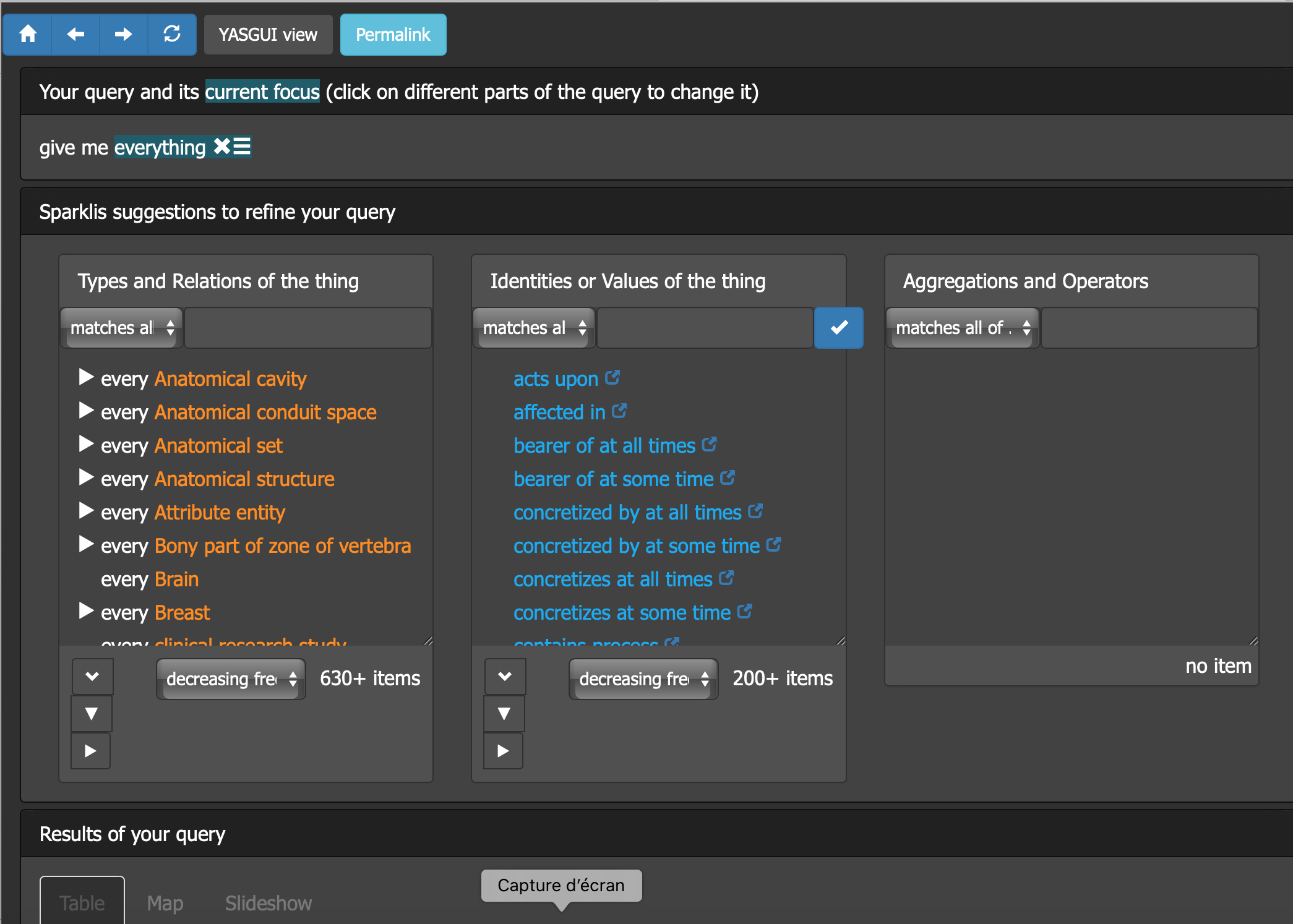


Figure 2: Main screen of Sparklis tool

* The upper part of the screen (introduced by *Your query and its current focus*) displays the current state of the query that is being built, translated into natural language; it allows the user selecting in this query the current concept of interest (called *current focus*), from which the current query can be refined.
* The middle part (introduced by *Sparklis suggestions* *to refine your query*) provides suggestions to extend and refine the query; is is composed of three parts materialized by three windows: i) the left one (named called *Types and Relations of the thing*) provides a list of the properties related to the current focus; ii) the one in the middle (named *Identities and Values* *of the thing*) provides the list of the values taken by the current concept of interest (IRIs of instances and data values); iii) the right-most window (named *Aggregations and Operators* ) suggests some processing that may be applied to the concept of interest (e.g. counting).
* The lower part of the screen (introduced by *Results of your query*) shows the result of the current SPARQL query.

Clicking in the box below *Types and Relations of the thing* causes a window to pop-up, allowing the user to enter a string corresponding to the name of the class from which the navigation should be started (Figure 3).

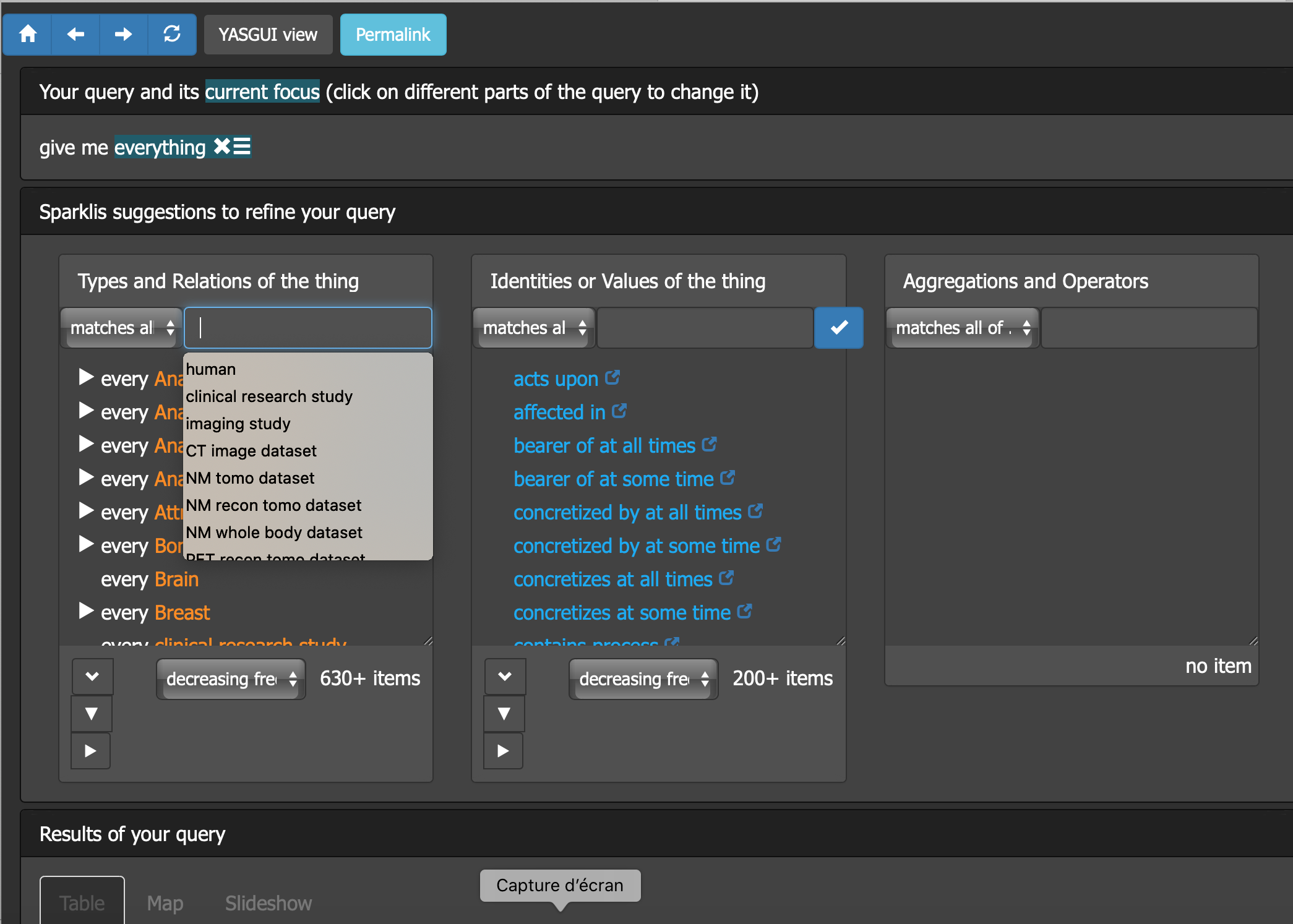


Figure 3: Selection of the class from which the navigation should be started (list of predefined terms in white window)

This string can be selected from a list of predefined class names that were considered as potentially of interest (*human*, *clinical research study*, etc). However, the user may enter any string and Sparklis will propose those classes that best match this string. The user is invited to select one on these, and this selection is integrated in the current SPARQL query, and the list of available properties updated to correspond to the new current focus. The number of occurrences is shown in parentheses after each property that is present in the graph. Any click appied to one of these properties will cause:

* the integration of this property in the current SPARQL query,
* the redefinition of the focus on the selected concept,
* the update of the list of values taken by this property in the property values window,
* the update of the suggestions of functions that may be applied,
* and the update of the result of the complete SPARQL query.

Note: it is important to mention that Sparklis takes into account inverse properties. Is means that a query can involve a relation in one direction while the RDF graph actually uses the inverse relation of this property.

At any time, the user may redefine the focus to any part of the current query (by a simple click on the entity of interest). This action leads to the updating of the the list of values taken by this property in the property values window and of the suggestions of functions that may be applied. Once the user has chosen his/her focus, then he/she can choose the property he/she is interested in, so that to extend or refine the query according to his/her wishes (Figure 4).

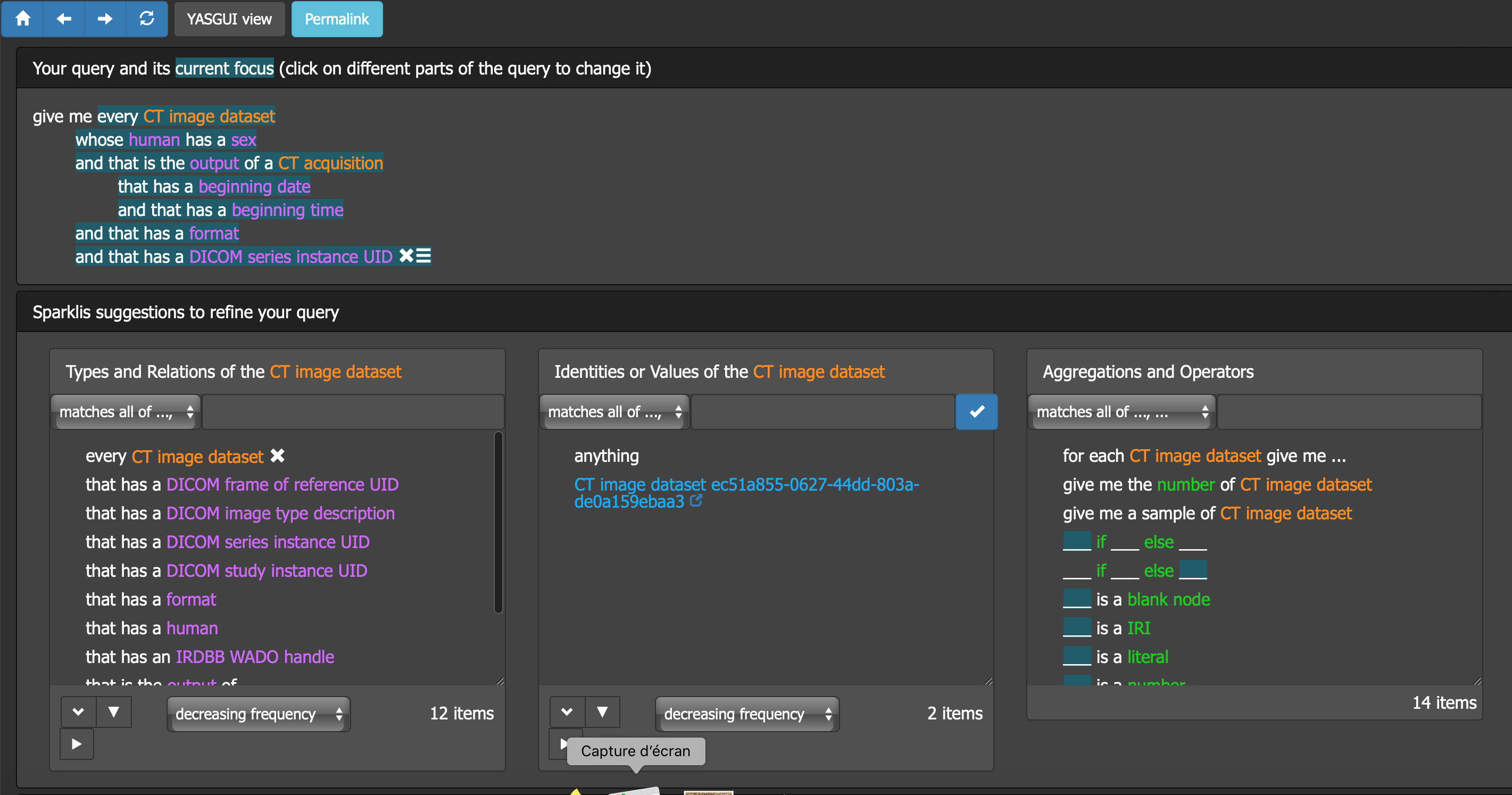


Figure 4: Example of query (the current focus is on *CT image dataset*, so the suggested properties are those which apply to this class)

At the top of the screen, the user can change the mode of display of the current SPARQL query from the *YASGUI* mode (*Yet Another Sparql GUI*, using natural language) to the *SPARKLIS View* mode, which allows displaying the SPARQL query that was generated (Figure 5).

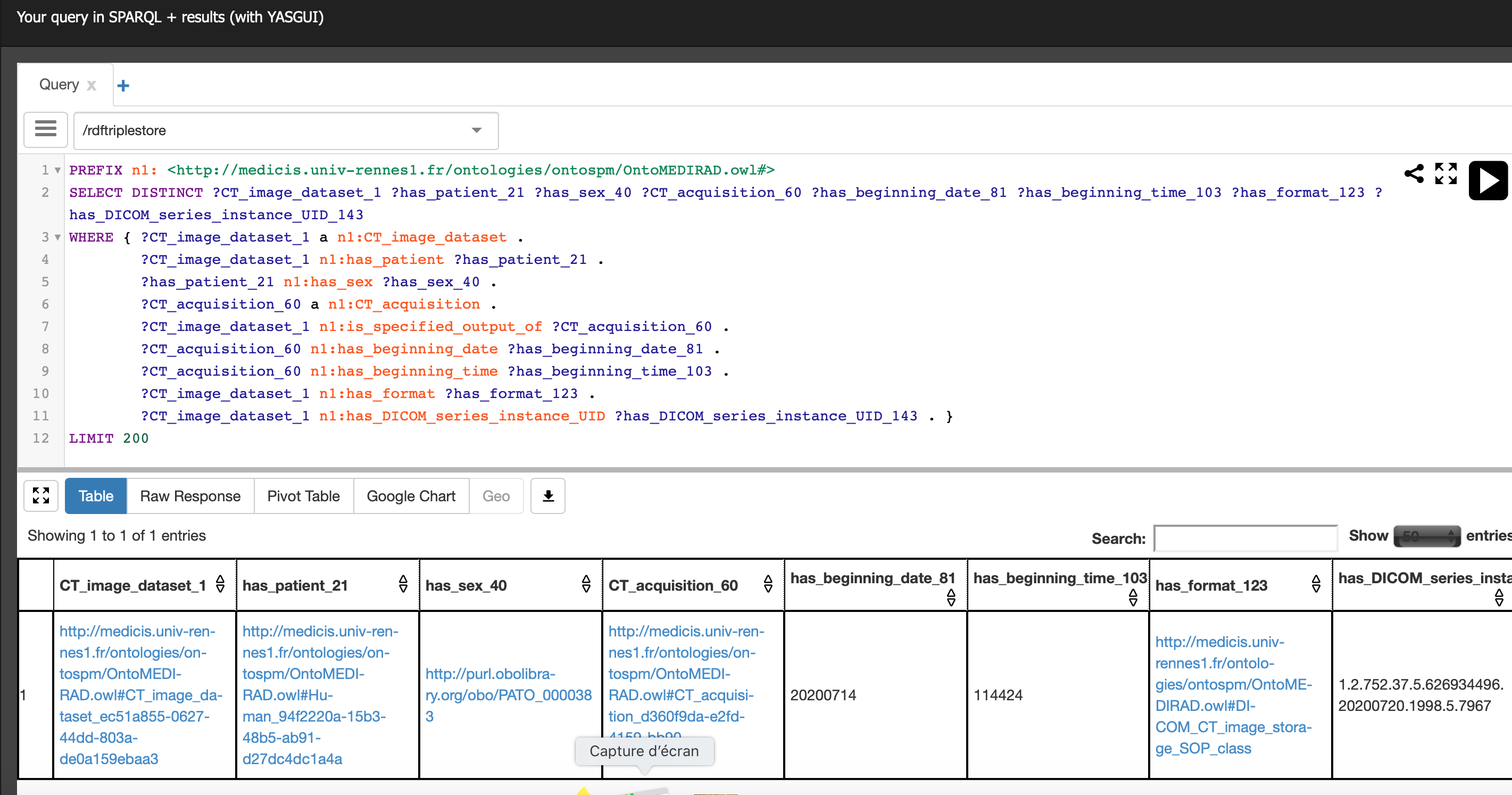


Figure 5: Example of SPARQL query (corresponding to the query shown Figure 4)

1. Implementation of the Sparklis tool

The Sparklis tool is implemented as a separate component of the IRDBB system, that complements the IRDBB\_UI web server by managing special pages providing a portal to the semantic graph. Sparklis accesses directly to the STARDOG’s SPARQL endpoint. (Figure 5). Sparklis is embedded in a Docker container, as all the other modules of IRDBB.

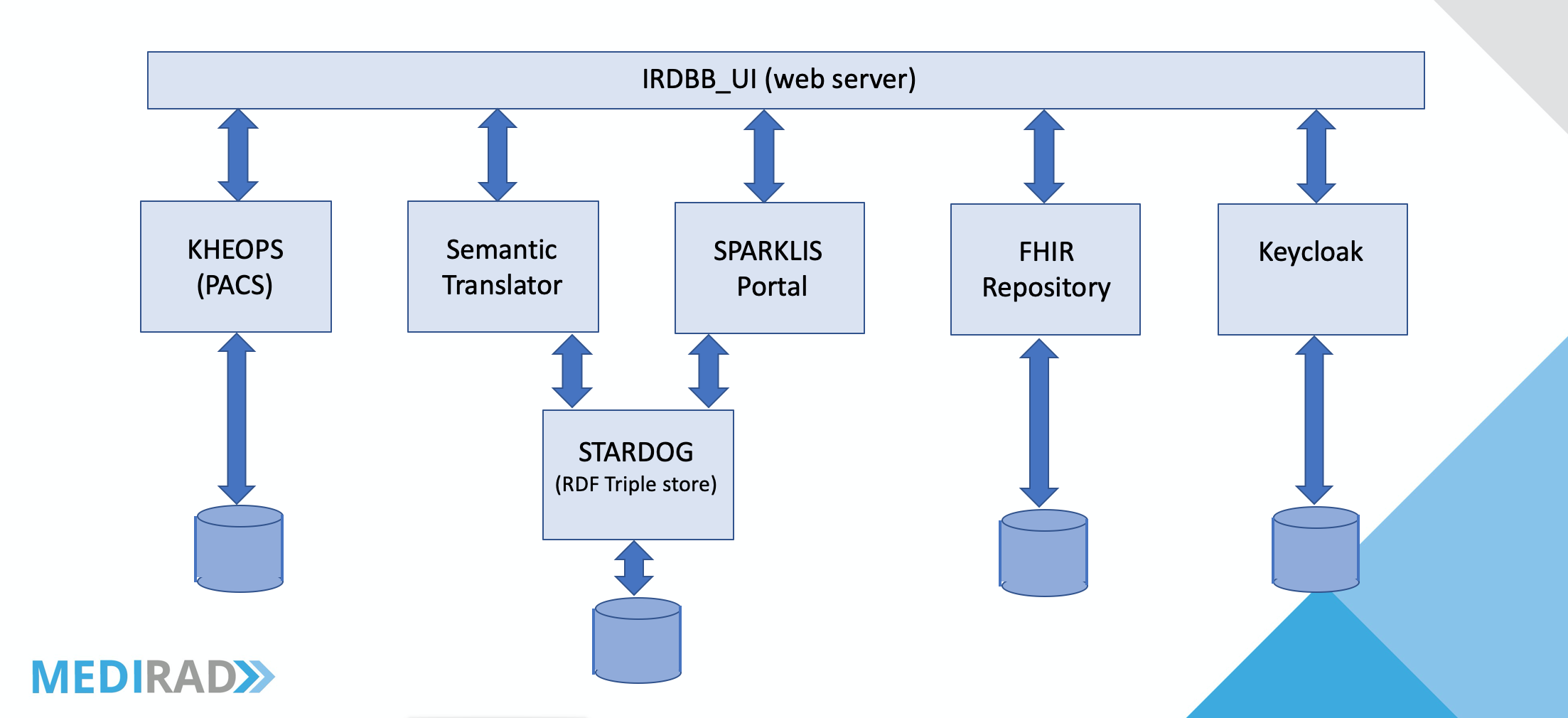


Figure 5: Architecture of the IRDBB system

The generic Sparklis engine (provided by its author Sébastien Ferré) was modified so that to:

* automatically connect to the STARDOG’s SPARQL endpoint,
* configure the way the classes and properties (object properties and data properties) used in the IRDBB RDF graph should be presented to the end-user.

This configuration allows selecting automatically the labels that should be presented to the user. This is very useful because some IRIs of the object properties and data properties involve numbers or identifiers with no human semantics. Therefore, the OntoMEDIRAD ontology uses an additional module (SPARKLISlabels.owl) that allows assigning dedicated labels that are well-suited to the way Sparklis translates SPARQL queries into natural language. These labels are provided in an annotation (called ontomedirad:sparklisLabel) associated to all the object properties and data properties.

The current configuration is shown in Table 1, and the Sparklis labels are listed in Annex 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Label**  **(First choice)** | **Label**  **(2nd choice)** | **Lang. pref. (First choice)** | **Lang. pref. 2nd choice** |
| **classes** | http://www.w3.org/2000/01/rdf-schema#label | http://www.w3.org/2004/02/skos/core#prefLabel | none | en |
| **properties** | http://medicis.univ-rennes1.fr/ontologies/ontospm/SPARKLISlabels.owl#sparklisLabel | none | none | none |

Table 1: Special configuration of Sparklis dedicated to ontoMEDIRAD

1. References

[1] SPARQL 1.1 https://www.w3.org/TR/sparql11-overview

[2] Sébastien Ferré. SPARKLIS: An Expressive Query Builder for SPARQL Endpoints with Guidance in Natural Language. Open Journal Of Semantic Web, Research Online Publishing (RonPub), 2017, 0. hal-01485093

[3] http://data.persee.fr/explorer/sparklis/

# Annex 1 Sparklis labels for OntoMEDIRAD

### Object properties

|  |  |  |  |
| --- | --- | --- | --- |
| **IRI of the object property** | **Value of rdfs:label** | **Value of skos:prefLabel** | **Value of ontomedirad:**  **sparklisLabel** |
| ontomedirad:acts\_upon |  | acts upon | acts upon |
| ontomedirad:affected\_in |  | affected in | affected in |
| ontomedirad:has\_agent |  | has agent | agent |
| ontomedirad:refers\_to\_attenuator |  | refers to attenuator | attenuator |
| purl:BFO\_0000158 | bearer of at all times |  | bearer of |
| purl:BFO\_0000053 | bearer of at some time |  | bearer of |
| ontomedirad:involves\_body\_part |  | involves body part | body part |
| purl:BFO\_0000165 | concretized by at all times |  | concretized by |
| purl:BFO\_0000058 | concretized by at some time |  | concretized by |
| purl:BFO\_0000164 | concretizes at all times |  | concretizes |
| purl:BFO\_0000059 | concretizes at some time |  | concretizes |
| purl:BFO\_0000067 | contains process |  | contains |
| ontomedirad:described\_by |  | described by | described by |
| ontomedirad:has\_destination\_  coordinate\_space\_specified\_by |  | has destination coordinate space specified by | destination coordinate space |
| ontomedirad:refers\_to\_device |  | refers to device | device |
| ontomedirad:is\_device\_setting\_of |  | is device setting of | device setting of |
| purl:BFO\_0000162 | has disposition at all times |  | disposition |
| purl:BFO\_0000112 | has disposition at some time |  | disposition |
| purl:BFO\_0000107 | disposition of at all times |  | disposition of |
| ontomedirad:has\_dose |  | has dose | dose |
| ontomedirad:is\_dose\_absorbed\_by |  | is dose absorbed by | dose absorbed by |
| ontomedirad:has\_duration |  | has duration | duration |
| ontomedirad:duration\_description\_of |  | is a duration description of | duration description of |
| ontomedirad:is\_duration\_of |  | is duration of | duration of |
| purl:BFO\_0000157 | during which exists |  | during which exists |
| ontomedirad:has\_equivalent\_  attenuator\_thickness |  | has equivalent attenuator thickness | equivalent attenuator thickness |
| ontomedirad:refers\_to\_  equivalent\_material |  | refers to equivalent material | equivalent material |
| ontomedirad:estimated\_from\_  phantom |  | estimated from phantom | estimated from phantom |
| purl:BFO\_0000108 | exists at |  | exists at |
| ontomedirad:extracted\_from |  | extracted from | extracted from |
| ontomedirad:has\_format |  | has format | format |
| ontomedirad:is\_format\_of |  | is format of | format of |
| ontomedirad:has\_instrument\_  function\_at\_some\_time |  | has instrument function at some time | function |
| purl:BFO\_0000085 | has function at some time |  | function |
| purl:BFO\_0000160 | has function at all times |  | function |
| ontomedirad:has\_instrument\_  function\_at\_all\_times |  | has instrument function at all times | function |
| purl:BFO\_0000079 | function of at all times |  | function of |
| purl:BFO\_0000101 | has generic dependent at some time |  | generic dependent |
| purl:BFO\_0000084 | generically depends on at some time |  | generically depends on |
| purl:BFO\_0000185 | has history |  | history |
| purl:BFO\_0000184 | history of |  | history of |
| ontomedirad:has\_patient |  | has patient | human |
| purl:BFO\_0000052 | inheres in at all times |  | inheres in |
| ontomedirad:has\_specified\_input |  | has specified input | input |
| ontomedirad:is\_specified\_input\_of |  | is specified input of | input of |
| ontomedirad:has\_instrument |  | has instrument | instrument |
| ontomedirad:is\_about\_  irradiation\_event |  | is about irradiation event | irradiation event |
| ontomedirad:is\_life\_of |  | is life of | life of |
| purl:BFO\_0000082 | located in at all times |  | located in |
| purl:BFO\_0000171 | located in at some time |  | located in |
| purl:BFO\_0000124 | has location at some time |  | location |
| purl:BFO\_0000170 | has location at all times |  | location |
| purl:BFO\_0000113 | has material basis at all times |  | material basis |
| purl:BFO\_0000163 | material basis of at all times |  | material basis of |
| purl:BFO\_0000127 | material basis of at some time |  | material basis of |
| ontomedirad:is\_quantity\_measured\_at |  | is quantity measured at | measured at |
| ontomedirad:is\_quantity\_measured\_in |  | is quantity measured in | measured in |
| purl:IAO\_0000039 | has measurement unit label |  | measurement unit |
| purl:BFO\_0000172 | has member part at all times |  | member part |
| purl:BFO\_0000115 | has member part at some time |  | member part |
| purl:BFO\_0000129 | member part of at some time |  | member part of |
| purl:BFO\_0000173 | member part of at all times |  | member part of |
| purl:BFO\_0000083 | occupies spatial region at some time |  | occupies spatial region |
| purl:BFO\_0000130 | occupies spatiotemporal region |  | occupies spatiotemporal region |
| purl:BFO\_0000155 | occupies temporal region |  | occupies temporal region |
| purl:BFO\_0000066 | occurs in |  | occurs in |
| ontomedirad:has\_specified\_output |  | has specified output | output |
| ontomedirad:is\_specified\_output\_of |  | is specified output of | output of |
| ontomedirad:is\_an\_output\_of |  | is an output of | output of |
| purl:BFO\_0000110 | has continuant part at all times |  | part |
| purl:BFO\_0000178 | has continuant part at some time |  | part |
| purl:BFO\_0000187 | has continuant part at all times that part exists |  | part |
| purl:BFO\_0000117 | has occurrent part |  | part |
| purl:BFO\_0000177 | part of continuant at all times |  | part of |
| purl:BFO\_0000176 | part of continuant at some time |  | part of |
| purl:BFO\_0000132 | part of occurrent |  | part of |
| purl:BFO\_0000186 | part of continuant at all times that whole exists |  | part of |
| ontomedirad:part\_of\_study |  | part of study | part of study |
| purl:BFO\_0000057 | has participant at some time |  | participant |
| purl:BFO\_0000167 | has participant at all times |  | participant |
| purl:BFO\_0000056 | participates in at some time |  | participates in |
| purl:BFO\_0000166 | participates in at all times |  | participates in |
| ontomedirad:preceded\_by |  | preceded by | preceded by |
| ontomedirad:precedes |  | precedes | precedes |
| ontomedirad:is\_about\_procedure |  | is about procedure | procedure |
| purl:BFO\_0000133 | process profile of |  | process profile of |
| purl:BFO\_0000119 | has profile |  | profile |
| purl:BFO\_0000153 | projects onto temporal region |  | projects onto |
| purl:BFO\_0000151 | projects onto spatial region at some time |  | projects onto spatial region |
| purl:BFO\_0000174 | has proper continuant part at some time |  | proper continuant part |
| purl:BFO\_0000111 | has proper continuant part at all times |  | proper part |
| purl:BFO\_0000118 | has proper occurrent part |  | proper part |
| purl:BFO\_0000138 | proper part of occurrent |  | proper part of |
| purl:BFO\_0000175 | proper part of continuant at some time |  | proper part of |
| purl:BFO\_0000137 | proper part of continuant at all times |  | proper part of |
| purl:BFO\_0000181 | has proper temporal part |  | proper temporal part |
| purl:BFO\_0000136 | proper temporal part of |  | proper temporal part of |
| ontomedirad:has\_protocol |  | has protocol | protocol |
| purl:BFO\_0000159 | has quality at all times |  | quality |
| purl:BFO\_0000086 | has quality at some time |  | quality |
| purl:IAO\_0000221 | is quality measurement of |  | quality measurement of |
| purl:BFO\_0000080 | quality of at all times |  | quality of |
| ontomedirad:has\_target\_radionuclide |  | has target radionuclide | radionuclide |
| purl:BFO\_0000054 | realized in |  | realized in |
| purl:BFO\_0000055 | realizes |  | realizes |
| ontomedirad:has\_reason\_for\_repeating\_acquisition |  | has reason for repeating acquisition | reason for repeating acquisition |
| ontomedirad:is\_equivalent\_dose\_received\_by |  | is equivalent dose received by | received by |
| ontomedirad:is\_effective\_dose\_received\_by |  | is effective dose received by | received by |
| ontomedirad:has\_target\_region |  | has target region | region |
| ontomedirad:represents |  | represents | represents |
| ontomedirad:is\_resampled\_on |  | is resampled on | resampled on |
| ontomedirad:is\_resampling\_of |  | is resampling of | resampling of |
| ontomedirad:is\_a\_result\_of |  | is a result of | result of |
| purl:BFO\_0000161 | has role at all times |  | role |
| purl:BFO\_0000087 | has role at some time |  | role |
| ontomedirad:is\_about\_role |  | is about role | role |
| purl:BFO\_0000081 | role of at all times |  | role of |
| ontomedirad:is\_scanned\_in |  | is scanned in | scanned in |
| ontomedirad:is\_about\_scanned\_object |  | is about scanned object | scanned object |
| ontomedirad:scans |  | scans | scans |
| ontomedirad:has\_scope\_of\_accumulation |  | has scope of accumulation | scope of accumulation |
| ontomedirad:has\_setting |  | has setting | setting |
| ontomedirad:has\_sex |  | has sex | sex |
| ontomedirad:has\_source\_coordinate\_space\_specified\_by |  | has source coordinate space specified by | source coordinate space |
| purl:BFO\_0000123 | has spatial occupant at some time |  | spatial occupant |
| purl:BFO\_0000152 | spatial projection of spatiotemporal at some time |  | spatial projection of spatiotemporal |
| purl:BFO\_0000126 | has spatiotemporal occupant |  | spatiotemporal occupant |
| purl:BFO\_0000125 | has specific dependent at some time |  | specific dependent |
| purl:BFO\_0000168 | has specific dependent at all times |  | specific dependent |
| purl:BFO\_0000070 | specifically depends on at all times |  | specifically depends on |
| purl:BFO\_0000169 | specifically depends on at some time |  | specifically depends on |
| ontomedirad:is\_about |  | is about | subject |
| purl:BFO\_0000156 | has temporal occupant |  | temporal occupant |
| purl:BFO\_0000121 | has temporal part |  | temporal part |
| purl:BFO\_0000139 | temporal part of |  | temporal part of |
| purl:BFO\_0000154 | temporal projection of spatiotemporal |  | temporal projection of |
| ontomedirad:treats |  | treats | treats |
| purl:uo#is\_unit\_of |  |  | unit of |
| ontomedirad:used\_as\_instrument\_in |  | used as instrument in | used in |
| ontomedirad:is\_protocol\_used\_in |  | is protocol used in | used in |

### Datatype properties

|  |  |  |  |
| --- | --- | --- | --- |
| **IRI of the datatype property** | **Value of rdfs:label** | **Value of skos:prefLabel** | **Value of ontomedirad:**  **sparklisLabel** |
| ontomedirad:has\_DICOM\_SOP\_class\_UID |  | has DICOM SOP class UID | DICOM SOP class UID |
| ontomedirad:has\_DICOM\_SOP\_instance\_UID |  | has DICOM SOP instance UID | DICOM SOP instance UID |
| ontomedirad:has\_DICOM\_UID |  | has DICOM UID | DICOM UID |
| ontomedirad:has\_DICOM\_frame\_of\_reference\_UID |  | has DICOM frame of reference UID | DICOM frame of reference UID |
| ontomedirad:has\_DICOM\_image\_type\_description |  | has DICOM image type description | DICOM image type description |
| ontomedirad:has\_DICOM\_series\_instance\_UID |  | has DICOM series instance UID | DICOM series instance UID |
| ontomedirad:has\_DICOM\_study\_instance\_UID |  | has DICOM study instance UID | DICOM study instance UID |
| ontomedirad:has\_IRDBB\_FHIR\_handle |  | has IRDBB FHIR handle | IRDBB FHIR handle |
| ontomedirad:has\_IRDBB\_WADO\_handle |  | has IRDBB WADO handle | IRDBB WADO handle |
| ontomedirad:inXSDdateTime |  | inXSDdateTime | XSDdateTime |
| ontomedirad:has\_beginning |  | has beginning | beginning |
| ontomedirad:has\_beginning\_date |  | has beginning date | beginning date |
| ontomedirad:has\_beginning\_time |  | has beginning time | beginning time |
| ontomedirad:has\_birthdate |  | has birthdate | birthdate |
| ontomedirad:has\_description |  | has description | description |
| ontomedirad:has\_device\_observer\_manufacturer\_name |  | has device observer manufacturer name | device observer manufacturer name |
| ontomedirad:has\_device\_observer\_model\_name |  | has device observer model name | device observer model name |
| ontomedirad:has\_device\_observer\_name |  | has device observer name | device observer name |
| ontomedirad:has\_device\_observer\_serial\_number |  | has device observer serial number | device observer serial number |
| ontomedirad:has\_end |  | has end | end |
| ontomedirad:has\_end\_date |  | has end date | end date |
| ontomedirad:has\_end\_time |  | has end time | end time |
| ontomedirad:has\_first\_name |  | has first name | first name |
| ontomedirad:hours |  | hours | hours |
| ontomedirad:has\_id |  | has id | id |
| ontomedirad:days |  | days | jours |
| ontomedirad:has\_location\_name |  | has location name | location name |
| ontomedirad:has\_lower\_limit |  | has lower limit | lower limit |
| ontomedirad:has\_manufacturer\_name |  | has manufacturer name | manufacturer name |
| ontomedirad:has\_measurement\_date\_and\_time |  | has measurement date and time | measurement date and time |
| purl:IAO\_0000039 | has measurement unit label |  | measurement unit |
| purl:IAO\_0000004 | has measurement value |  | measurement value |
| ontomedirad:has\_method\_name |  | has method name | method name |
| ontomedirad:minutes |  | minutes | minutes |
| ontomedirad:has\_model\_name |  | has model name | model name |
| ontomedirad:months |  | months | months |
| ontomedirad:has\_name |  | has name | name |
| http://www.w3.org/2004/02/skos/core#notation | notation |  | notation |
| ontomedirad:has\_number\_of\_X-Ray\_sources |  | has number of X-Ray sources | number of X-Ray sources |
| ontomedirad:has\_number\_of\_irradiation\_events |  | has number of irradiation events | number of irradiation events |
| ontomedirad:has\_rank |  | has rank | rank |
| ontomedirad:seconds |  | seconds | seconds |
| ontomedirad:has\_title |  | has title | title |
| ontomedirad:has\_upper\_limit |  | has upper limit | upper limit |
| ontomedirad:weeks |  | weeks | weeks |
| ontomedirad:years |  | years | years |