SciCatProject/panosc-search-api PR #101 review
version 2021-11-04

Suggestions

- Store and build locally techniques tree. Uncouple PaNOSC search from Bioportal.
  Instead of making calls to Bioportal as needed (aka everytime we receive a user query), we suggest to download the techniques ontology file at set times. Once the techniques file has been downloaded, the system should build the tree locally and populate a local cache. Every user request should be served using the local cache.
- Allow only querying by technique pids.
- Clearly document that the techniques operators have an extended functionality and are not just plain boolean operators.
- Refactor the list of technique operators as explained in the next section.
- Consider as an alternative solution to expand the technique tagging when datasets are created or updated in the cataloging system.
  At creation or update time, the techniques tags should be expanded from the tacg provided to the list of the tag and all its ancestors.
  This approach will increase the number of tags on each datasets, but will revert techniques operators back to normal boolean logic operators, eliminating the need of the complex logic illustrated below.

Proposed operators

All the operators listed in this section are *deep operators*, meaning that they are converted to a set operator where the set is composed by the technique provided and all its descendants.
Therefore we define *technique descendants set* as the set encompassing the root technique and all its descendants.
Here is the proposed list of operators for the techniques:

- **Equal to** (operator: *eq*)
  It selects any dataset that is tagged with a technique that is contained in the technique descendants set.
  Syntax:
  ○ short form : { "pid" : "my-technique" }
  ○ long form : { "pid" : { "eq" : "my-technique"}}
- **Not Equal to** (operator: *not*)
  It selects any dataset that is tagged with any technique that is not contained in the technique descendants set.
  Syntax:
  ○ { "pid" : { "not" : "my-technique"}}
- **And** (operator: *and*)
  It selects any dataset that is tagged with at least two techniques and each one of them belongs to at least one operand technique descendants set.
  Syntax:
  ○ { "pid" : { "and" : ["my-technique-1", "my-technique-2"]}
- **Or** (operator: *or*)
  It selects any dataset that is tagged with one or more techniques which belongs to at least one operand technique descendants set.
  Syntax:
  ○ { "pid" : { "or" : ["my-technique-1", "my-technique-2"]}

Examples

In order to provide meaningful examples, we need to define few datasets and the related set of techniques which they are tagged with.

Datasets and tagged techniques:

Dataset Id	Techniques
1	Resonant inelastic x-ray scattering
2	Anomalous solution x-ray scattering
3	Grazing incidence small angle x-ray scattering
4	Inelastic x-ray small angle scattering
5	Soft x-ray small angle scattering
6	Magnetic x-ray tomography
7	X-ray emission spectroscopy
8	X-ray emission spectroscopy, magnetic x-ray tomography
9	X-ray emission spectroscopy, grazing incidence small angle scattering
10	X-ray emissions spectroscopy, anomalous solution x-ray scattering

Notes:
- Only the relevant portion of the query is shown below.
- Techniques are referred by their name for sake of clarity of the example. In a real situation, the name will be substituted by the technique pid.

Query 1

Use case:
User is looking for datasets that are tagged with the term *X-ray emission spectroscopy* or any of its descendants.

User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : "x-ray emission spectroscopy"
        }
      }
    }
  ]
  ...
}
```
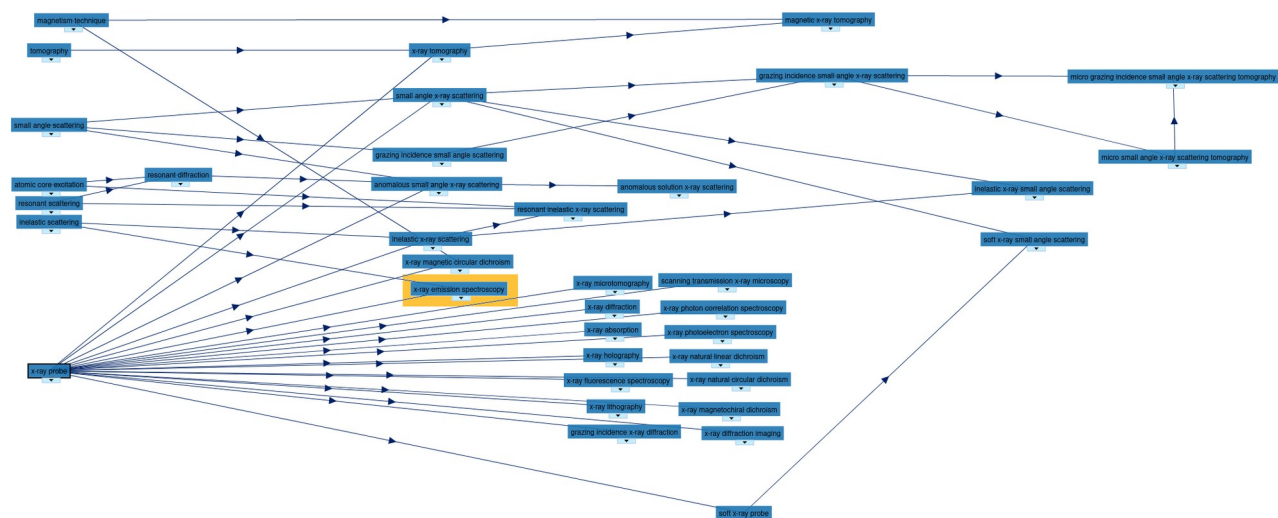
Equivalent user query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : {"eq" : "x-ray emission spectroscopy"}
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid": { "inq" : [
            "x-ray emission spectroscopy"
          ]
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 7, 8, 9, 10

Highlighted in yellow is the technique descendants set for *x-ray emission spectroscopy* technique

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	
2	Anomalous solution x-ray scattering	
3	Grazing incidence small angle x-ray scattering	
4	Inelastic x-ray small angle scattering	
5	Soft x-ray small angle scattering	
6	Magnetic x-ray tomography	
7	X-ray emission spectroscopy	Y
8	X-ray emission spectroscopy, magnetic x-ray tomography	Y
9	X-ray emission spectroscopy, grazing incidence small angle scattering	Y
10	X-ray emissions spectroscopy, anomalous solution x-ray scattering	Y

Query 2

Use case:
User is searching for datasets tagged with technique *small angle scattering* or any of their descendants

User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : "small angle scattering"
        }
      }
    }
  ]
  ...
}
```
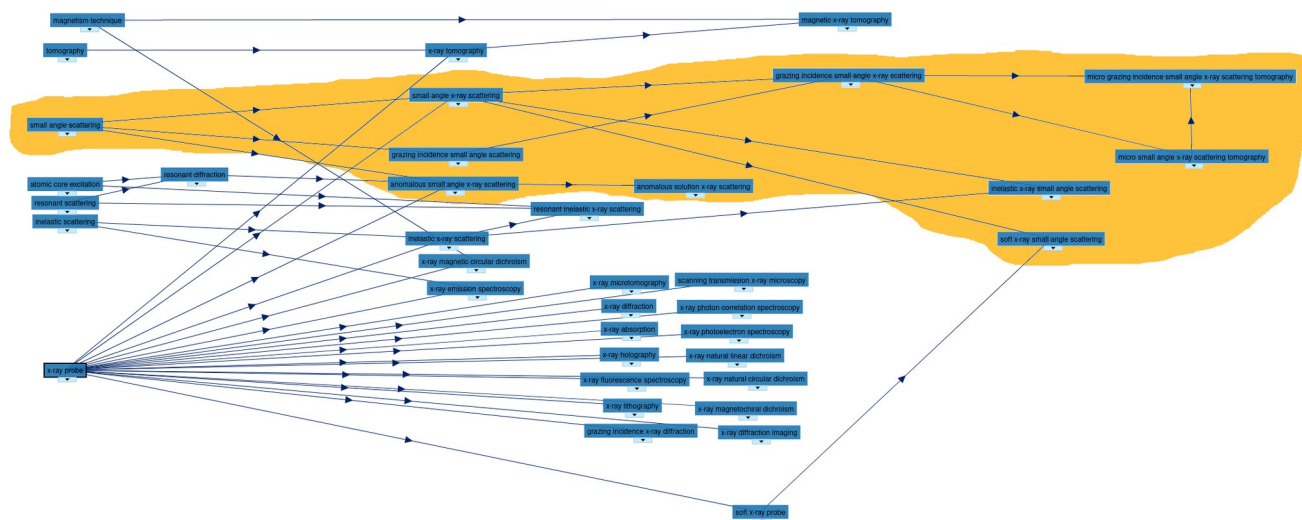
Equivalent User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "eq" : "small angle scattering" }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid": {"inq" : [
            "small angle scattering",
            "small angle x-ray scattering",
            "anomalous small angle x-ray scattering",
            "grazing incidence small angle scattering",
            "anomalous solution x-ray scattering",
            "grazing incidence small angle x-ray scattering",
            "micro small angle x-ray scattering tomography",
            "micro grazing incidence small angle x-ray scattering tomography",
            "soft x-ray small angle scattering",
            "inelastic x-ray small angle scattering"
          ] }
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 2, 3, 4, 5, 9, 10

Highlighted in yellow is the technique descendants set for techniques *small angle scattering*.

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	
2	Anomalous solution x-ray scattering	y
3	Grazing incidence small angle x-ray scattering	y
4	Inelastic x-ray small angle scattering	y
5	Soft x-ray small angle scattering	y
6	Magnetic x-ray tomography	
7	X-ray emission spectroscopy	
8	X-ray emission spectroscopy, magnetic x-ray tomography	
9	X-ray emission spectroscopy, grazing incidence small angle scattering	Y
10	X-ray emissions spectroscopy, anomalous solution x-ray scattering	Y

Query 3

Use case:
User is looking for datasets that are not tagged with the term *X-ray emission spectroscopy* or any of its descendants.
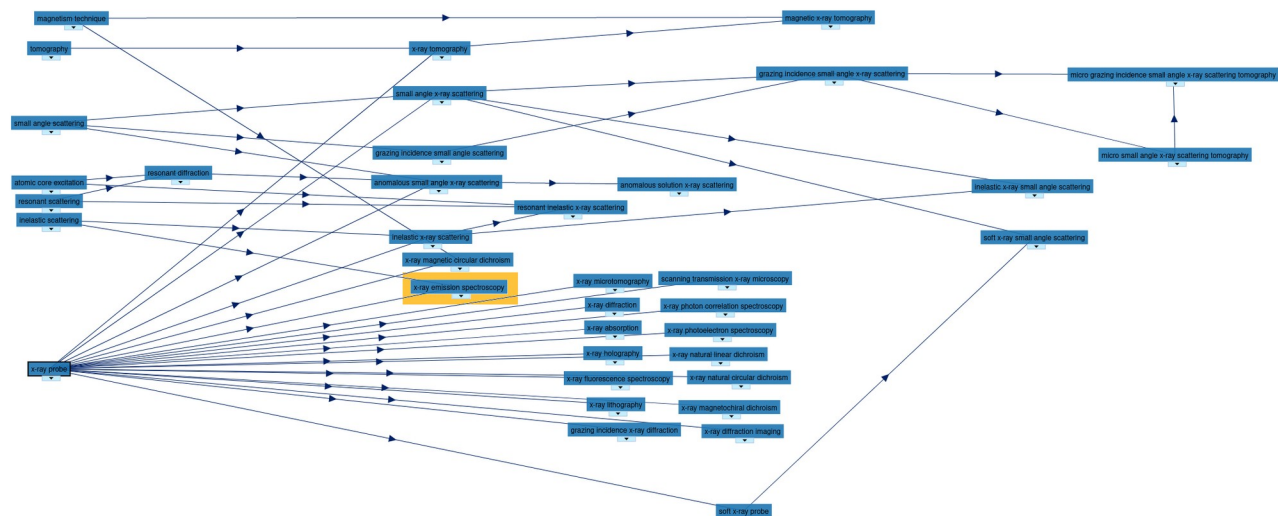
User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "not" : "x-ray emission spectroscopy" }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid": { "nin" : [
            "x-ray emission spectroscopy"
          ]
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 1, 2, 3, 4, 5, 6

Highlighted in yellow is the technique descendants set for *x-ray emission spectroscopy* technique

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	Y
2	Anomalous solution x-ray scattering	Y
3	Grazing incidence small angle x-ray scattering	Y
4	Inelastic x-ray small angle scattering	Y
5	Soft x-ray small angle scattering	Y
6	Magnetic x-ray tomography	Y
7	X-ray emission spectroscopy	
8	X-ray emission spectroscopy, magnetic x-ray tomography	
9	X-ray emission spectroscopy, grazing incidence small angle scattering	
10	X-ray emission spectroscopy, anomalous solution x-ray scattering	

Query 4

Use case:

User is searching for datasets that are not tagged tagged with technique *small angle scattering* or any of their descendants

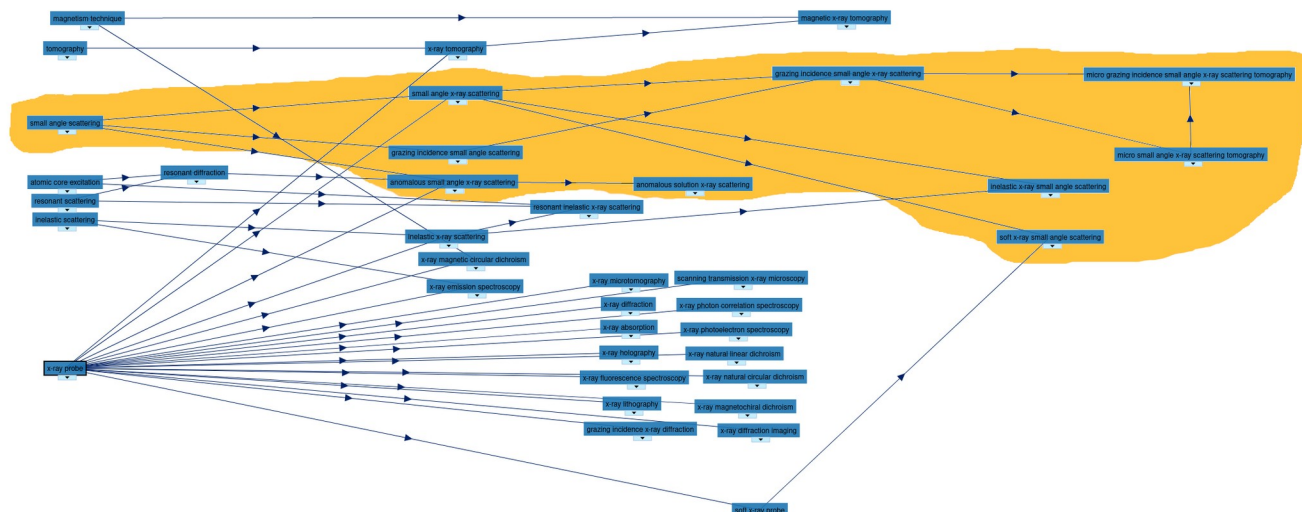User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "not" : "small angle scattering" }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid": {"nin" : [
            "small angle scattering",
            "small angle x-ray scattering",
            "anomalous small angle x-ray scattering",
            "grazing incidence small angle scattering",
            "anomalous solution x-ray scattering",
            "grazing incidence small angle x-ray scattering",
            "micro small angle x-ray scattering tomography",
            "micro grazing incidence small angle x-ray scattering tomography",
            "soft x-ray small angle scattering",
            "inelastic x-ray small angle scattering"
          ] }
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 1, 6, 7, 8

Highlighted in yellow is the technique descendants set for techniques *small angle scattering*.
We want to select datasets that are not tagged with any of the techniques that are in this set.

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	Y
2	Anomalous solution x-ray scattering	
3	Grazing incidence small angle x-ray scattering	
4	Inelastic x-ray small angle scattering	
5	Soft x-ray small angle scattering	
6	Magnetic x-ray tomography	Y
7	X-ray emission spectroscopy	Y
8	X-ray emission spectroscopy, magnetic x-ray tomography	Y
9	X-ray emission spectroscopy, grazing incidence small angle scattering	
10	X-ray emissions spectroscopy, anomalous solution x-ray scattering	

Query 5

Use case:
User is searching for datasets tagged with both techniques *small angle scattering* and *magnetism technique* or any of their descendants
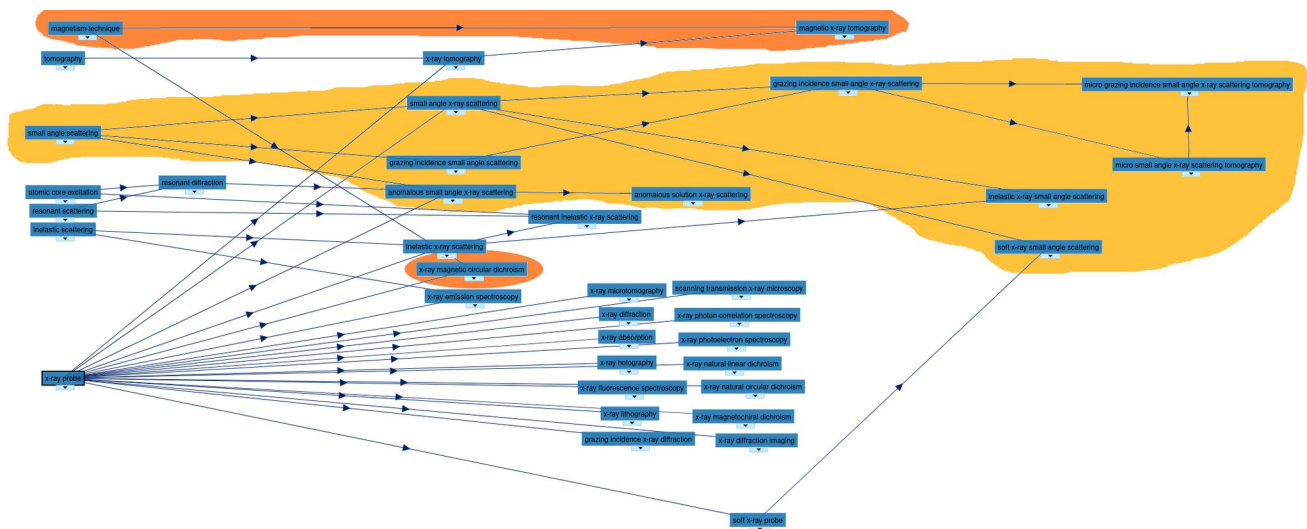
User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "and": [
            "small angle scattering",
            "magnetism technique
          ] }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "and" : [
            {
              "pid": {"inq" : [
                "small angle scattering",
                "small angle x-ray scattering",
                "anomalous small angle x-ray scattering",
                "grazing incidence small angle scattering",
                "anomalous solution x-ray scattering",
                "grazing incidence small angle x-ray scattering",
                "micro small angle x-ray scattering tomography",
                "micro grazing incidence small angle x-ray scattering tomography",
                "soft x-ray small angle scattering",
                "inelastic x-ray small angle scattering"
              ] }
            },
            {
              "pid": {"inq" : [
                "magnetism technique",
                "x-ray magnetic circular dichroism",
                "magnetic x-ray tomography"
              ] }
            }
          ]
        }
      }
    }
  ]
  ...
}
```

Datasets returned: None.
There are no datasets that are tagged with two techniques each belonging to the technique descendants set of *small angle scattering* and *magnetism technique*

Highlighted in orange is the technique descendants set of *magnetism technique*, while in yellow is the technique descendants set of *small angle scattering*. Each datasets must be tagged with at least one technique from both sets to be selected.

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	
2	Anomalous solution x-ray scattering	
3	Grazing incidence small angle x-ray scattering	
4	Inelastic x-ray small angle scattering	
5	Soft x-ray small angle scattering	
6	Magnetic x-ray tomography	
7	X-ray emission spectroscopy	
8	X-ray emission spectroscopy, magnetic x-ray tomography	
9	X-ray emission spectroscopy, grazing incidence small angle scattering	
10	X-ray emissions spectroscopy, anomalous solution x-ray scattering	

Query 6

Use case:
User is searching for datasets tagged with techniques *small angle scattering* or technique *magnetism technique* or any of their descendants

User Query:
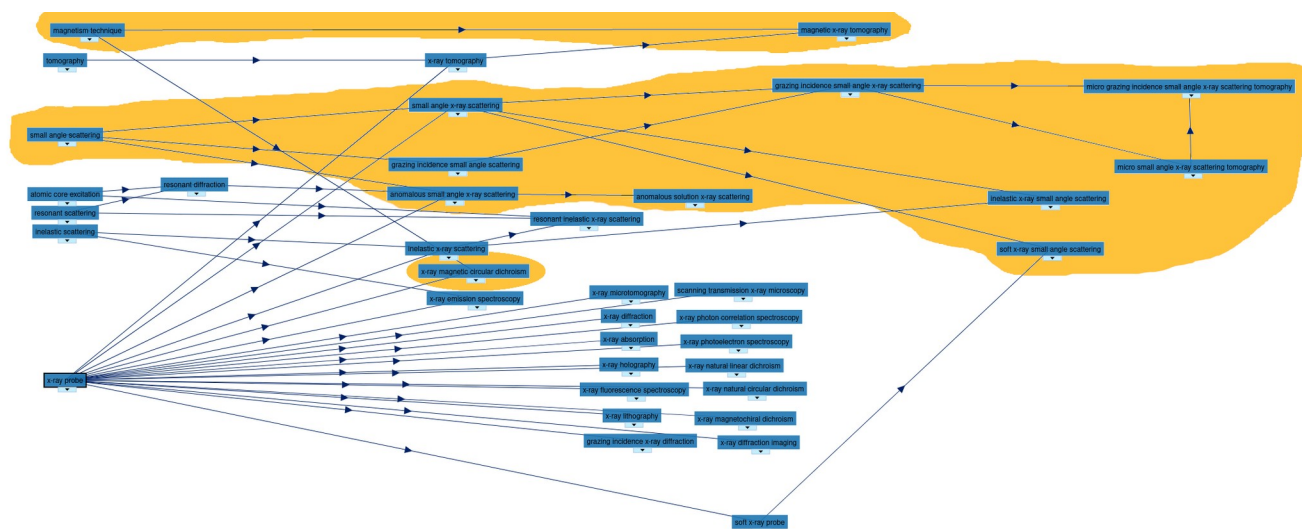```
{
   ...
   "include": [
      {
         "relation": "techniques",
         "scope": {
            "where": {
               "pid" : { "or": [
                  "small angle scattering",
                  "magnetism technique
               ] }
            }
         }
      }
   ]
   ...
}
```

Interpreted query:
```
{
   ...
   "include": [
      {
         "relation": "techniques",
         "scope": {
            "where": {
               "pid" : {
                  "inq" : [
                     "small angle scattering",
                     "small angle x-ray scattering",
                     "anomalous small angle x-ray scattering",
                     "grazing incidence small angle scattering",
                     "anomalous solution x-ray scattering",
                     "grazing incidence small angle x-ray scattering",
                     "micro small angle x-ray scattering tomography",
                     "micro grazing incidence small angle x-ray scattering tomography",
                     "soft x-ray small angle scattering",
                     "inelastic x-ray small angle scattering",
                     "magnetism technique",
                     "x-ray magnetic circular dichroism",
                     "magnetic x-ray tomography"
                  ]
               }
            }
         }
      }
   ]
   ...
}
```

Datasets returned: 2, 3, 4, 5, 6, 8, 9, 10

Highlighted in yellow is the set with all the admissible techniques. The logic or between two techniques descendants sets is equivalent to the technique belonging to the union of the technical descendants sets. Each dataset is select if any of the techniques is tagged with belongs the union of the sets.

Datasets selection

| Dataset Id | Techniques | Selected |
|:---:|:---|:---:|
| 1 | Resonant inelastic x-ray scattering | |
| 2 | Anomalous solution x-ray scattering | Y |
| 3 | Grazing incidence small angle x-ray scattering | Y |
| 4 | Inelastic x-ray small angle scattering | Y |
| 5 | Soft x-ray small angle scattering | Y |
| 6 | Magnetic x-ray tomography | Y |
| 7 | X-ray emission spectroscopy | |
| 8 | X-ray emission spectroscopy, magnetic x-ray tomography | Y |
| 9 | X-ray emission spectroscopy, grazing incidence small angle scattering | Y |
| 10 | X-ray emissions spectroscopy, anomalous solution x-ray scattering | Y |

## Query 7

Use case:
user is looking for datasets tagged with both techniques *small angle scattering* and *x-ray probe* or any of their descendants
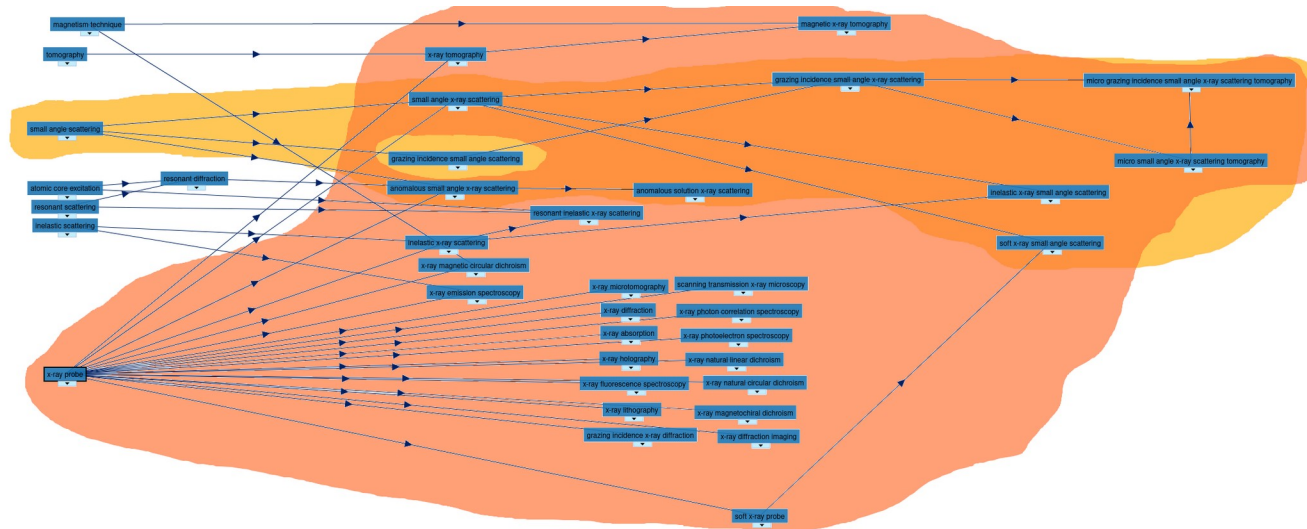
User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "and": [
            "small angle scattering",
            "x-ray probe"
          ] }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "and" : [
            {
              "pid": {"inq" : [
                "small angle scattering",
                "small angle x-ray scattering",
                "anomalous small angle x-ray scattering",
                "grazing incidence small angle scattering",
                "anomalous solution x-ray scattering",
                "grazing incidence small angle x-ray scattering",
                "micro small angle x-ray scattering tomography",
                "micro grazing incidence small angle x-ray scattering tomography",
                "soft x-ray small angle scattering",
                "inelastic x-ray small angle scattering",
              ] }
            },
            {
              "pid": {"inq" : [
                "x-ray probe",
                "x-ray emission spectroscopy",
                "resonant inelastic x-ray scattering",
                "magnetic x-ray tomography",
                "anomalous small angle x-ray scattering",
                "anomalous solution x-ray scattering",
                "grazing incidence small angle x-ray scattering",
                "micro small angle x-ray scattering tomography",
                "micro grazing incidence small angle x-ray scattering tomography",
                "soft x-ray small angle scattering",
                ...
              ] }
            }
          ]
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 2, 3, 4, 5, 9, 10
These datasets are tagged with one or more element from both technique descendants sets.



Highlighted in yellow is the technique descendants set for technique *small angle scattering*, while in orange is the technique descendants set for technique *x-ray probe*. Each selected dataset must be tagged with one technique from both sets.

Datasets selection

| Dataset Id | Techniques | Selected |
|---|---|---|
| 1 | Resonant inelastic x-ray scattering | |
| 2 | Anomalous solution x-ray scattering | Y |
| 3 | Grazing incidence small angle x-ray scattering | Y |
| 4 | Inelastic x-ray small angle scattering | Y |
| 5 | Soft x-ray small angle scattering | Y |
| 6 | Magnetic x-ray tomography | |
| 7 | X-ray emission spectroscopy | |
| 8 | X-ray emission spectroscopy, magnetic x-ray tomography | |
| 9 | X-ray emission spectroscopy, grazing incidence small angle scattering | Y |
| 10 | X-ray emissions spectroscopy, anomalous solution x-ray scattering | Y |

# Query 8

Use case:

user is searching for datasets tagged with either techniques *small angle scattering* and *x-ray probe* or any of their descendants
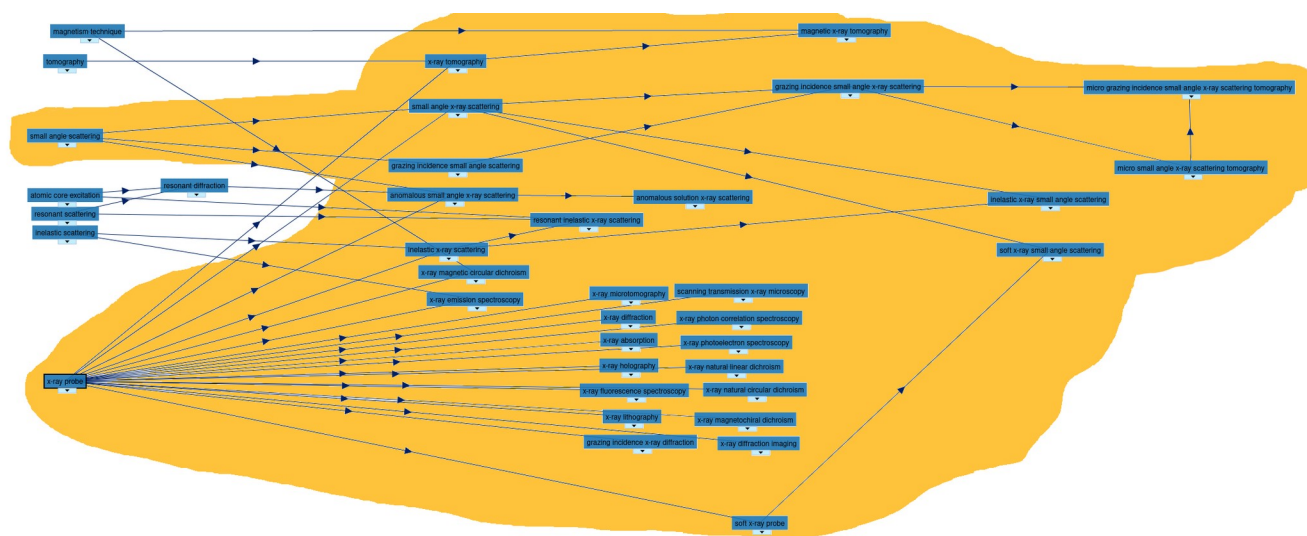
User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "or": [
            "small angle scattering",
            "x-ray probe"
          ] }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid": {"inq" : [
            "small angle scattering",
            "small angle x-ray scattering",
            "anomalous small angle x-ray scattering",
            "grazing incidence small angle scattering",
            "anomalous solution x-ray scattering",
            "grazing incidence small angle x-ray scattering",
            "micro small angle x-ray scattering tomography",
            "micro grazing incidence small angle x-ray scattering tomography",
            "soft x-ray small angle scattering",
            "inelastic x-ray small angle scattering",
            "x-ray probe",
            "x-ray emission spectroscopy",
            "resonant inelastic x-ray scattering",
            "magnetic x-ray tomography",
            "anomalous small angle x-ray scattering",
            "anomalous solution x-ray scattering",
            "grazing incidence small angle x-ray scattering",
            "micro small angle x-ray scattering tomography",
            "micro grazing incidence small angle x-ray scattering tomography",
            "soft x-ray small angle scattering",
            "inelastic x-ray tomography",
            ...
          ] }
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Highlighted in yellow is the techniques set resulting from the union of the technique descendants sets of the techniques *small angle scattering* and *x-ray probe*. A selected dataset must be tagged with at least one technique belonging to the resulting set.

Detasets selection

| Dataset Id | Techniques | Selected |
|:---:|:---|:---:|
| 1 | Resonant inelastic x-ray scattering | Y |
| 2 | Anomalous solution x-ray scattering | Y |
| 3 | Grazing incidence small angle x-ray scattering | Y |
| 4 | Inelastic x-ray small angle scattering | Y |
| 5 | Soft x-ray small angle scattering | Y |
| 6 | Magnetic x-ray tomography | Y |
| 7 | X-ray emission spectroscopy | Y |
| 8 | X-ray emission spectroscopy, magnetic x-ray tomography | Y |
| 9 | X-ray emission spectroscopy, grazing incidence small angle scattering | Y |
| 10 | X-ray emissions spectroscopy, anomalous solution x-ray scattering | Y |

# Query 9

Use case:

user is looking for datasets tagged with technique *small angle scattering* or any of its descendants but not with *x-ray probe* or any of its descendants
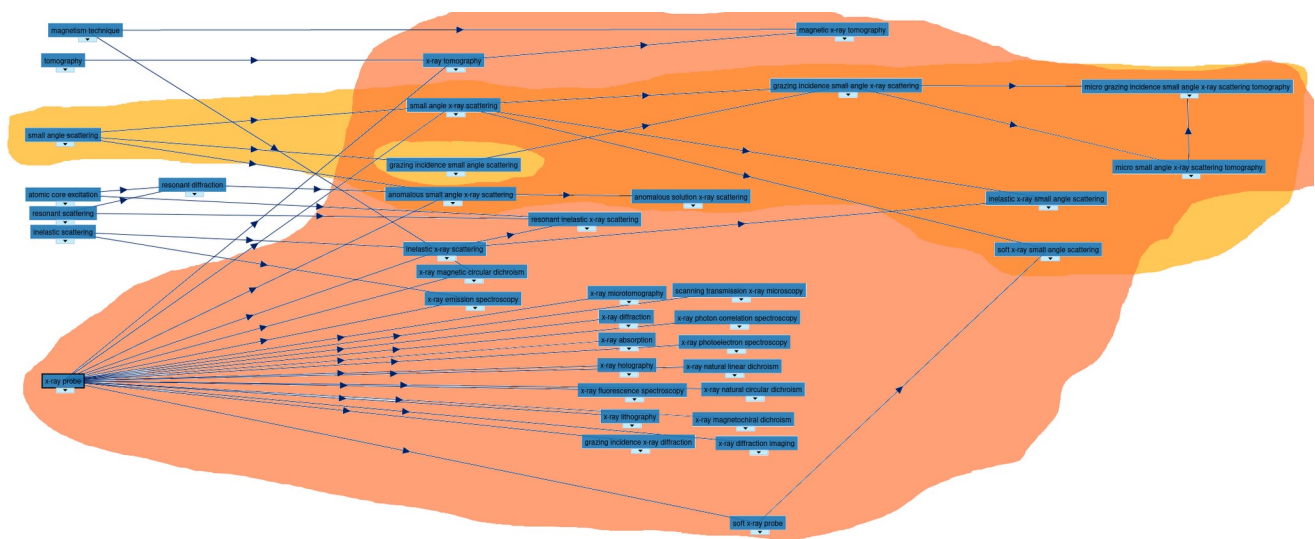
User Query:

```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "and": [
            "small angle scattering",
            { "not" : "x-ray probe" }
          ] }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:

```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "and" : [
            {
              "pid": {"inq" : [
                "small angle scattering",
                "small angle x-ray scattering",
                "anomalous small angle x-ray scattering",
                "grazing incidence small angle scattering",
                "anomalous solution x-ray scattering",
                "grazing incidence small angle x-ray scattering",
                "micro small angle x-ray scattering tomography",
                "micro grazing incidence small angle x-ray scattering tomography",
                "soft x-ray small angle scattering",
                "inelastic x-ray small angle scattering",
              ] }
            },
            {
              "pid": {"nin" : [
                "x-ray probe",
                "x-ray emission spectroscopy",
                "resonant inelastic x-ray scattering",
                "magnetic x-ray tomography",
                "anomalous small angle x-ray scattering",
                "anomalous solution x-ray scattering",
                "grazing incidence small angle x-ray scattering",
                "micro small angle x-ray scattering tomography",
                "micro grazing incidence small angle x-ray scattering tomography",
                "soft x-ray small angle scattering",
                ...
              ] }
            }
          ]
        }
      }
    }
  ]
  ...
}
```

Datasets returned: None

Highlighted in yellow is the technique descendants set for technique *small angle scattering*, while in orange is the technique descendants set for technique *x-ray probe*. Each selected dataset must be tagged with one technique from the yellow set, but not with any technique from the orange set.

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	
2	Anomalous solution x-ray scattering	
3	Grazing incidence small angle x-ray scattering	
4	Inelastic x-ray small angle scattering	
5	Soft x-ray small angle scattering	
6	Magnetic x-ray tomography	
7	X-ray emission spectroscopy	
8	X-ray emission spectroscopy, magnetic x-ray tomography	
9	X-ray emission spectroscopy, grazing incidence small angle scattering	
10	X-ray emission spectroscopy, anomalous solution x-ray scattering	

Query 10

Use case:

User is searching for datasets tagged with technique *small angle scattering* or any of its descendants but not with *magnetism technique* or any of its descendants
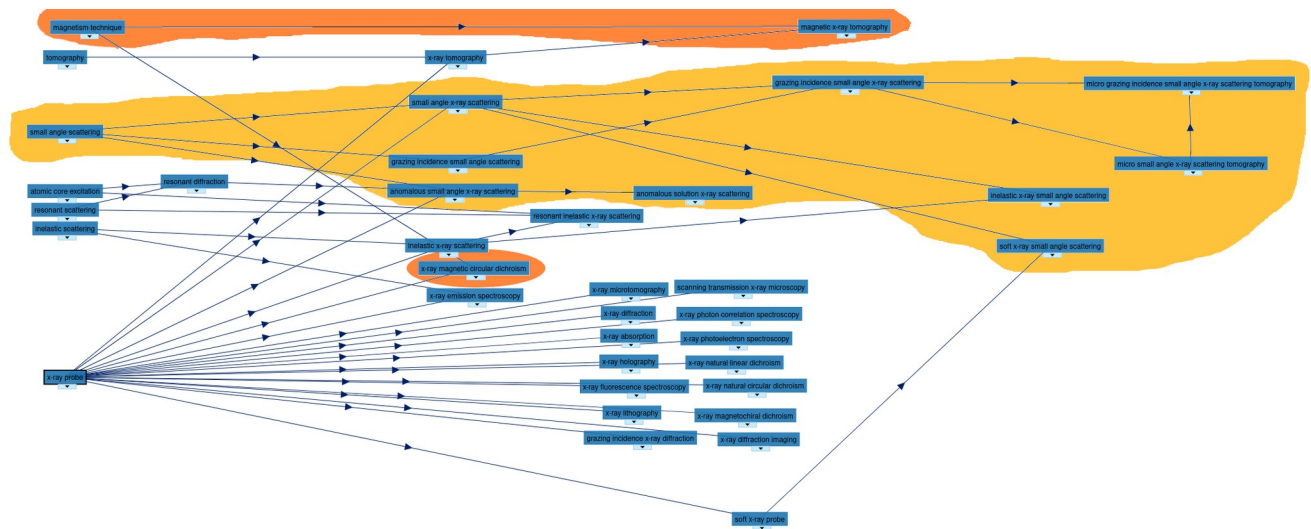
User Query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "pid" : { "and": [
            "small angle scattering",
            { "not" : "magnetism technique }
          ] }
        }
      }
    }
  ]
  ...
}
```

Interpreted query:
```
{
  ...
  "include": [
    {
      "relation": "techniques",
      "scope": {
        "where": {
          "and" : [
            {
              "pid": {"inq" : [
                "small angle scattering",
                "small angle x-ray scattering",
                "anomalous small angle x-ray scattering",
                "grazing incidence small angle scattering",
                "anomalous solution x-ray scattering",
                "grazing incidence small angle x-ray scattering",
                "micro small angle x-ray scattering tomography",
                "micro grazing incidence small angle x-ray scattering tomography",
                "soft x-ray small angle scattering",
                "inelastic x-ray small angle scattering"
              ] }
            },
            {
              "pid": {"nin" : [
                "magnetism technique",
                "x-ray magnetic circular dichroism",
                "magnetic x-ray tomography"
              ] }
            }
          ]
        }
      }
    }
  ]
  ...
}
```

Datasets returned: 2, 3, 4, 5, 9, 10

Highlighted in orange is the technique descendants set of *magnetism technique*, while in yellow is the technique descendants set of *small angle scattering*. Each selected dataset must be tagged with any of the techniques in the yellow set, but not with any on the orange set.

Datasets selection

Dataset Id	Techniques	Selected
1	Resonant inelastic x-ray scattering	
2	Anomalous solution x-ray scattering	Y
3	Grazing incidence small angle x-ray scattering	Y
4	Inelastic x-ray small angle scattering	Y
5	Soft x-ray small angle scattering	Y
6	Magnetic x-ray tomography	
7	X-ray emission spectroscopy	
8	X-ray emission spectroscopy, magnetic x-ray tomography	
9	X-ray emission spectroscopy, grazing incidence small angle scattering	Y
10	X-ray emissions spectroscopy, anomalous solution x-ray scattering	Y