Intelligent Queries over BIRN Data using the Foundational Model of Anatomy and a Distributed Query-based Data Integration System

JF Brinkley, MD, PhD[1], JA Turner, PhD[2], LT Detwiler, MS[1], JLV Mejino, Jr, MD[1], ME Martone, PhD[3], and DL Rubin, MD, MS[4]

[1]Structural Informatics Group, University of Washington, Seattle, WA; [2]Mind Research Network, Albuquerque, NM; [3]Dept of Neurosciences, UC San Diego; [4]Center for Biomedical Informatics Research, Stanford University

Summary

•The goal is to demonstrate the use of an ontology and a distributed query-based data integration system to perform "intelligent" queries over distributed functional magnetic resonance imaging (fMRI) data.

•The fMRI data come from a study of schizophrenic (sz) versus healthy subjects performing a simple motor task.

•The ontology is a neuroanatomical view of the Foundational Model of Anatomy (FMA). The distributed query system is called DXBrain.

•The question is whether the distribution of fMRI activations in various brain regions varies between healthy and sz individuals.

•Image data are retrieved from BIRN (Fig 1) , spatially normalized to a common coordinate system, and analyzed for contiguous clusters of activation, each of which is summarized by the point of maximal activation. This point is automatically annotated with an anatomical label by the Internet Talairach demon. A summary set of activations are provided in a spreadsheet, which is converted to XML and made web accessible (Fig 5).

•In DXBrain an XQuery (Fig 2) first queries the FMA to find the parts of a given region (the DLPFC) and their associated Talairach labels as represented in the FMA (Fig 3), simplifies the resulting list (Fig 4), retains only those that are annotated by a label in the list, divides the retained sites into healthy and sz, adds a distinguishing color to each site (Fig 5), and displays the results using our MindSeer visualizer (Fig 6).

•All these processes, including the imaging pipeline, can be automated and run over multiple data sources on the Internet, using neuroanatomical and other semantic information from ontologies to reduce the need for manual search, and to facilitate knowledge discovery.

Funding: NIH HL08770, RR021992, RR025736 and HHSN268200800020C.

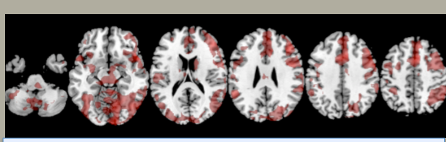

Fig 1. Example fMRI data from the Biomedical Informatics Research Network (BIRN). Regions of activations are in red.

```
(:::::::::::::Healthy versus Schizophrenic activations confined to a specified region of the brain::::::::::::::::::::::::::::::::::::::::::::::)
declare namespace fma="http://sig.uw.edu/fma#";
declare namespace rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#";
declare namespace rdfs="http://www.w3.org/2000/01/rdf-schema#";

(:::::::::::::Set parameters. These could be set by a GUI ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::)
let $h_color := 'Red'
let $s_color := 'Yellow'
let $shape := 'BIG_SPHERE'
let $size := '1.75'

(:::::::::::::Embedded query to ontology web service to find parts of the Dorsolateral prefrontal cortex - DLPFC ::::::::::::::::::::::::)
let $wsdlURL := "http://ontviews.biostr.washington.edu:8080/VSparQL_Service/wsdl/VSparQLService.wsdl"
let $serviceName := "VSparQLService"
let $methodName := "executeQuery"

let $query :=
"PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX gleen:<java:edu.washington.sig.gleen.>
PREFIX fma:<http://sig.uw.edu/fma#>
PREFIX tmp:<http://sig.biostr.washington.edu/tmp#>
PREFIX apf: <java:com.hp.hpl.jena.query.pfunction.library.>

CONSTRUCT {?structure fma:Talairach ?structure_name .  ?lateralized_structure fma:Talairach ?lateralized_structure_name . ?part fma:Talairach ?part_name}
FROM <http://sig.uw.edu/neurofma_v1.3>
WHERE
{
?structure apf:assign fma:Dorsolateral_prefrontal_cortex .
OPTIONAL {?structure fma:Talairach ?t1 . ?t1 fma:name ?structure_name  }.
?lateralized_structure rdfs:subClassOf ?structure .
OPTIONAL {?lateralized_structure fma:Talairach ?t2 . ?t2 fma:name ?lateralized_structure_name } .
?lateralized_structure gleen:OnPath ( '([fma:regional_part]|[fma:constitutional_part])*' ?part ) .
OPTIONAL {?part fma:Talairach ?t3 . ?t3 fma:name ?part_name }
}"

let $regions := dxq:xqueryWS($wsdlURL, $serviceName, $methodName, $query)

(:::::::::::::Use XQuery to reformat the results returned from the ontology into a  simple XML list :::::::::::::::::::::::::::::::::::::::::::::)
let $sub_regions_tals :=
     for $sub in $regions//rdf:Description
     return
       <fmatal>
         {$sub//fma:Talairach/text()}
       </fmatal>

(:::::::::::::Retrieve the  BIRN summary activation clusters::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::)
let $act_doc := doc("http://synapse.biostr.washington.edu/~brinkley/dxdemos/BIRNSzSummaryData.xml")
let $raw_sitecount := count($act_doc//site)

(:::::::::::::Keep only  activation clusters  in the specified region, add fixed display parameters ::::::::::::::::::::::::::::::::::::::::::::::)
let $act_by_tal :=
   for $el in $act_doc//site
   for  $fmatal in $sub_regions_tals
   return if (fn:contains($el/Tal,$fmatal))
   then
      <site> {$el/*} {$fmatal}<Size>{$size}</Size><Shape>{$shape}</Shape> </site>
   else ()
let $filtered_sitecount := count($act_by_tal)

(:::Divide the retained activation clusters  into healthy versus schizophrenic populations  and set different display colors::::::::::)
let $healthy_sites :=
    for $el in $act_by_tal[Diagnostic_group/text()='h']
      return <site> {$el/*} <Color>{$h_color}</Color></site>
let $healthy_sitecount := count($healthy_sites)

let $schizo_sites :=
   for $el in $act_by_tal[Diagnostic_group/text()='s']
      return <site> {$el/*} <Color>{$s_color}</Color></site>
let $schizo_sitecount := count($schizo_sites)

(:::::::::::::Return the two lists as a single XML document, suitable for input to the visualization program MindSeer ::::::::::::::::::::)
return
<results>
   <raw_sitecount>{$raw_sitecount}</raw_sitecount>
   <filtered_sitecount>{$filtered_sitecount}</filtered_sitecount>
   <healthy_sitecount>{$healthy_sitecount}</healthy_sitecount>
   <schizo_sitecount>{$schizo_sitecount}</schizo_sitecount>
  <patient>
      <pnum>MNI</pnum>
        {$healthy_sites}
        {$schizo_sites}
   </patient>
</results>
```

Change this to change region, e.g. Frontal_lobe

Fig 2. Executable XQuery at http://xbrain.biostr.washington.edu:8080/dxbrain/TestQuery.jsp?query_id=312

Fig 3. Embedded vSparQL query to the FMA to find the parts of the Dorsolateral prefrontal cortex (DLPFC) and their associated Talairach labels. Results of this query are shown on the right. The part in bold matches the bold terms in Figs 4 and 5.


```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:apf="java:com.hp.hpl.jena.query.pfunction.library."
  xmlns:tmp="http://sig.biostr.washington.edu/tmp#"
  xmlns:gleen="java:edu.washington.sig.gleen."
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:fma="http://sig.uw.edu/fma#" >
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_46_of_right_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_46_of_left_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_left_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_left_superior_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Left Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_right_superior_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Right Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
 <rdf:Description rdf:about="http://sig.uw.edu/fma#Brodmann_area_9_of_right_middle_frontal_gyrus">
  <fma:Talairach rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fma:Talairach>
 </rdf:Description>
</rdf:RDF>
```


Fig 4. Create simplified list of Talairach labels corresponding to the DLPFC. The label in bold matches the example site shown in Fig 5.


```
<?xml version="1.0" encoding="UTF-8"?>
<results>
  <fmatal>Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fmatal>
  <fmatal>Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fmatal>
  <fmatal>Left Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
  <fmatal>Left Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
  <fmatal>Right Cerebrum.Frontal Lobe.Superior Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
  <fmatal>Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 9</fmatal>
</results>
```



```
<site>
   <site_label>310542253_5.55</site_label>
   <BIRN_ID>310542253</BIRN_ID>
   <Diagnostic_group>h</Diagnostic_group>
   <Voxels>651</Voxels>
   <Max_Z>5.55</Max_Z>
   <P>3.74E-4</P>
   <log10P>3.43</log10P>
   <Distance>18</Distance>
   <right_coord>50</right_coord>
   <sup_coord>30</sup_coord>
   <ant_coord>40</ant_coord>
   <Tal>      Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</Tal>
   <fmatal>Right Cerebrum.Frontal Lobe.Middle Frontal Gyrus.Gray Matter.Brodmann area 46</fmatal>
   <Size>1.75</Size>
   <Shape>BIG_SPHERE</Shape>
   <Color>Red</Color>
</site>
```


Fig 5. Retrieve summary data, keep only those that are in the DLPFC, divide into two diagnostic groups, add distinguishing colors. Above shows an example activation site from summary BIRN data that is located in the DLPFC. Elements added by the XQuery are bold. <fmatal> is added to show the automatic match with the FMA. The BIRN summary data has 1741 such activation sites.

Region	Filtered	Healthy (Red)	Schizo-phrenic (Yellow)
Dorso_lateral prefrontal cortex	55	25	30
Brodmann area 6	213	107	106
Frontal lobe	575	288	287

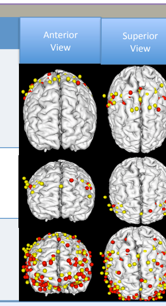

Fig 6. Summary of retrieved data. Each row shows a region, the number of sites retained after removing sites not in the region (Filtered), the number of healthy sites in the filtered data, and the number of schizophrenic sites. Images show sites superimposed on the Colin standard atlas brain. Sites do not always fall on the brain surface because the spatial normalization is coarse. For this simple motor task there is no significant difference in distribution.