# Data Integration && SPARQL

У. Төрболд

ХШУИС, Мэдээллийн технологийн III-р түвшний оюутан, turbold1125@gmail.com

# ОРШИЛ

Энэхүү лабораторын ажлаар МУИС-ийн өгөгдлийн сангаас хүснэгтэн өгөгдлийг татаж авах, Karma Data Integration tool ашиглан RDF график мэдээллийн сан үүсгэх үйл явцыг баримтжуулахад оршино. Өгөгдлийн семантик бүтцийг тодорхойлоход МУИС-ийн онтологийг ашиглана. Мөн үүссэн график мэдээллийн санг GraphDB-д нэгтгэх, өгөгдлийн сангаас тодорхой мэдээллийг авахын тулд сонгон асуулга боловсруулах зэрэг орно

# ЗОРИЛГО

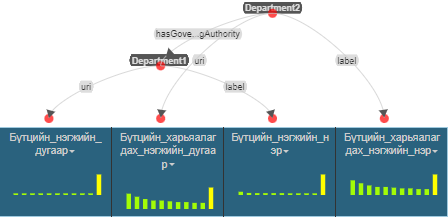
1. МУИС-ийн мэдээллийн порталаас 5 хүснэгтийн мэдээллийн багцыг татаж авах
2. Хүснэгтийн өгөгдлийг RDF формат руу хөрвүүлэхийн тулд Karma Data Integration tool ашиглах
3. RDF мэдээллийн семантик бүтцийг тодорхойлохын тулд МУИС-ийн онтологийг ашиглах
4. Өөрчлөгдсөн өгөгдөл, онтологийг ашиглан RDF график мэдээллийн санг үүсгэх.
5. Хадгалалт болон query – г үр ашигтай ажиллахын тулд график мэдээллийн санг GraphDB-д импортлох.
6. График мэдээллийн сангаас мэдээлэл авахын тулд дор хаяж 3 select query бичих.

# 3. ОНОЛЫН СУДАЛГАА

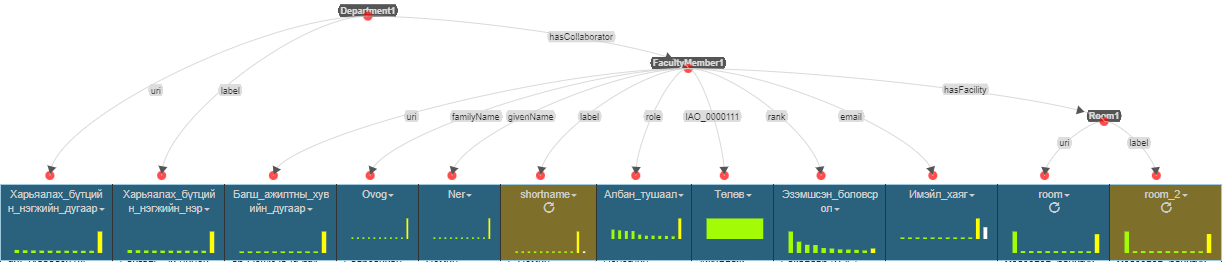
1. Хүснэгтийн өгөгдлийг татаж авах: https://data.num.edu.mn/ сайтаас таван хүснэгтэн өгөгдлийг сонгон татаж авах. Эдгээр өгөгдлүүд нь оюутны бүртгэл, хичээлийн хуваарь, багш нарын мэдээлэл, судалгааны мэдээлэл болон бусад холбогдох мэдээллийн эх сурвалж зэрэг МУИС-ийн домайнаас олон төрлийн мэдээллийг төлөөлөх ёстой.
2. RDF график мэдээллийн сан үүсгэх: Татаж авсан хүснэгтэн өгөгдлийг RDF формат руу хөрвүүлэхийн тулд Karma Data Integration tool – ыг ашиглах. Үүнд МУИС-ийн онтологид тохирох ойлголт, хамаарал, шинж чанаруудтай хүснэгтийн өгөгдлийн баганыг зураглах бөгөөд өөрчлөгдсөн өгөгдөл нь RDF-ийн өгөгдлийн загварт нийцсэн байх ёстой бөгөөд мэдээллийг илэрхийлнэ.
3. МУИС-ийн онтологийг ашиглах: Үүсгэсэн RDF график мэдээллийн санд семантик уялдаа холбоо, харилцан ажиллах чадварыг хангахын тулд МУИС-ийн онтологийг ашиглах юм. Онтологи нь МУИС-ийн домэйны онцлогт тохирсон ойлголт, харилцаа, шинж чанаруудын урьдчилан тодорхойлсон багцыг өгдөг. Хүснэгтийн өгөгдлийг онтологитой уялдуулах нь өгөгдлийн ойлголт, нэгтгэлийг сайжруулж, өөр өөр байгууллагуудын хооронд утга учиртай холболт, харилцааг бий болгоно.
4. GraphDB-д хадгалах: RDF график мэдээллийн сангийн удирдлагын систем болох GraphDB-д оруулах нь RDF-ийн өгөгдлийг GraphDB-д импортлох, зөв хадгалах, индексжүүлэх, асуулга хийх чадварыг баталгаажуулах юм.
5. Select query: GraphDB ашиглан RDF график мэдээллийн санд дор хаяж гурван Select query боловсруулж, гүйцэтгэх нь график мэдээллийн санг нэгдсэн өгөгдлөөс утга учиртай, гүнзгий мэдээлэл олж авах чадварыг харуулах ёстой.

# 4. ХЭРЭГЖҮҮЛЭЛТ

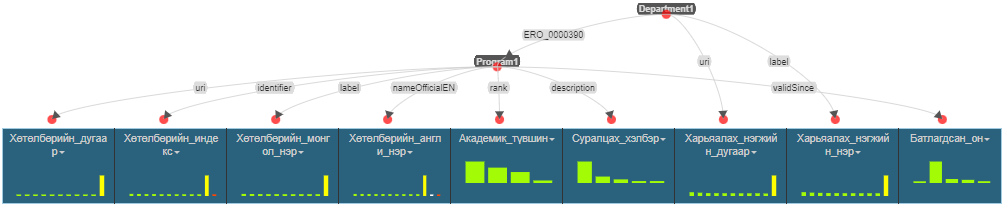
1. Байгууллагын бүтэц өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



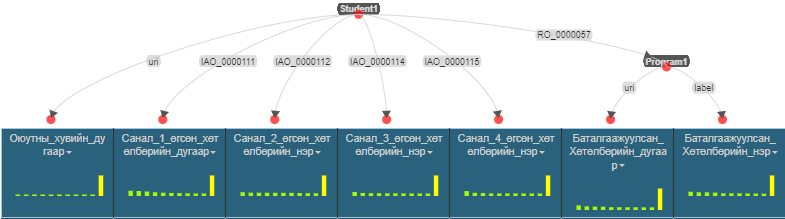
1. Багш ажилчдын нэрс өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



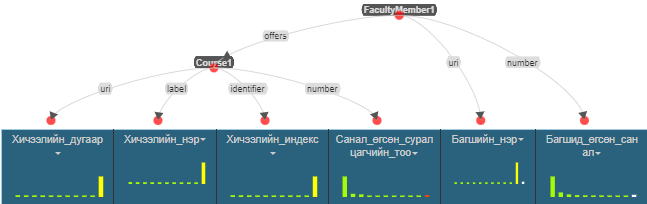
1. Сургалтын хөтөлбөр өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



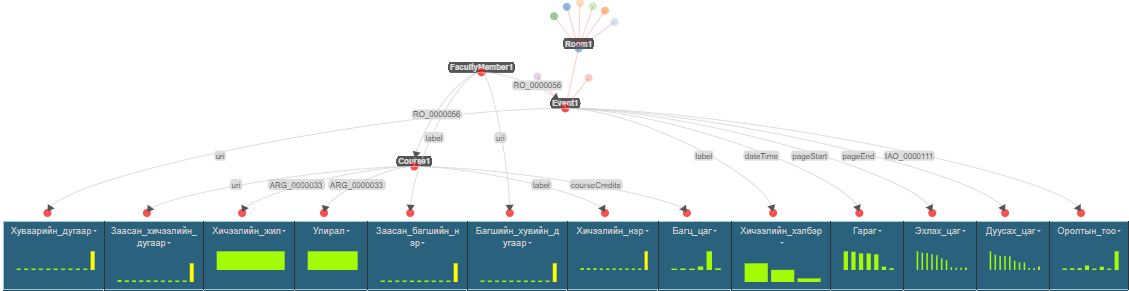
1. Хөтөлбөр сонголтын баталгаажуулалт өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал

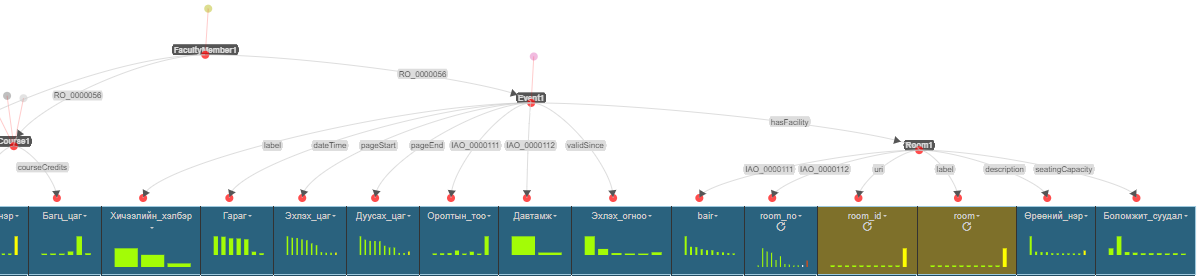


1. Хөтөлбөр сонгох санал өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



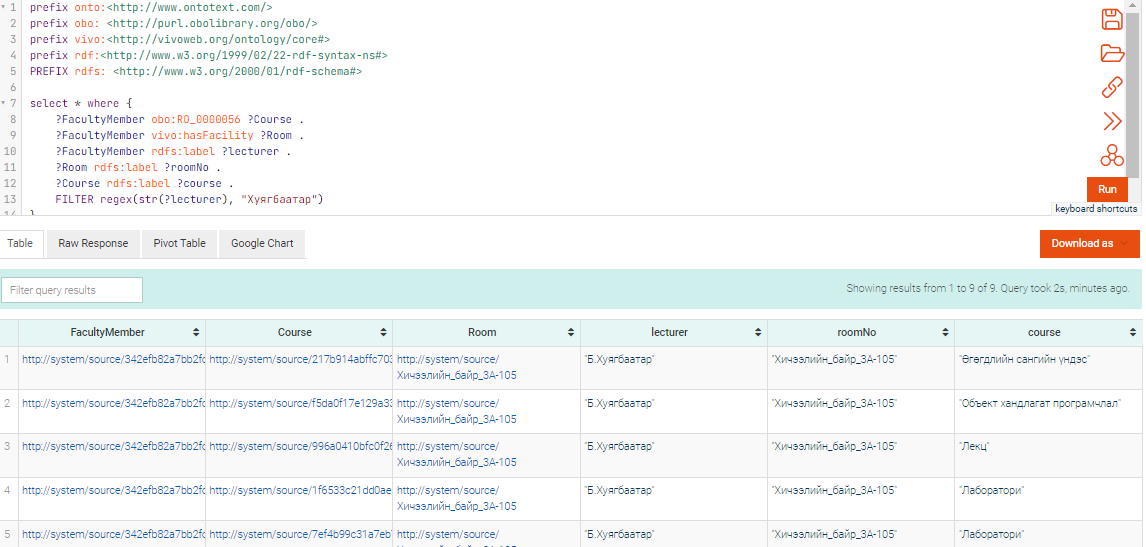
1. Хичээлийн хуваарь өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



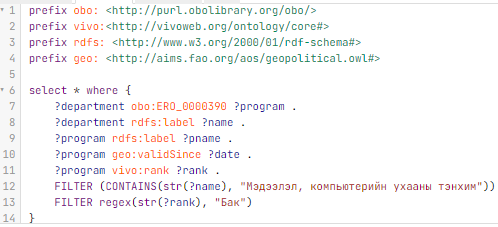


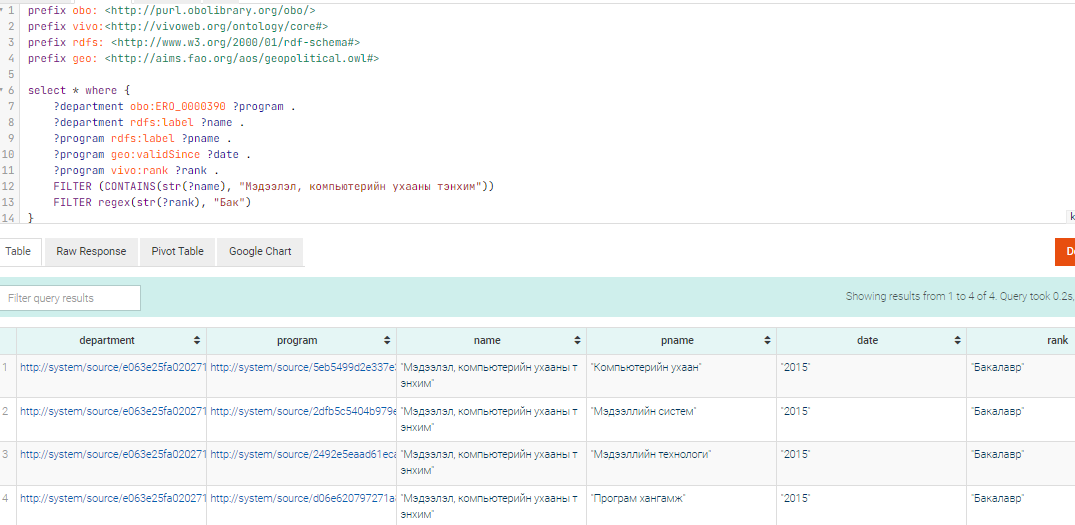
4.2.1 Хуягбаатар багшийн ямар ямар хичээл хаана ордог вэ?



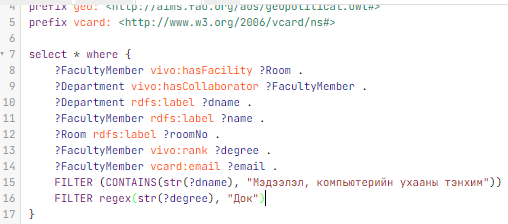


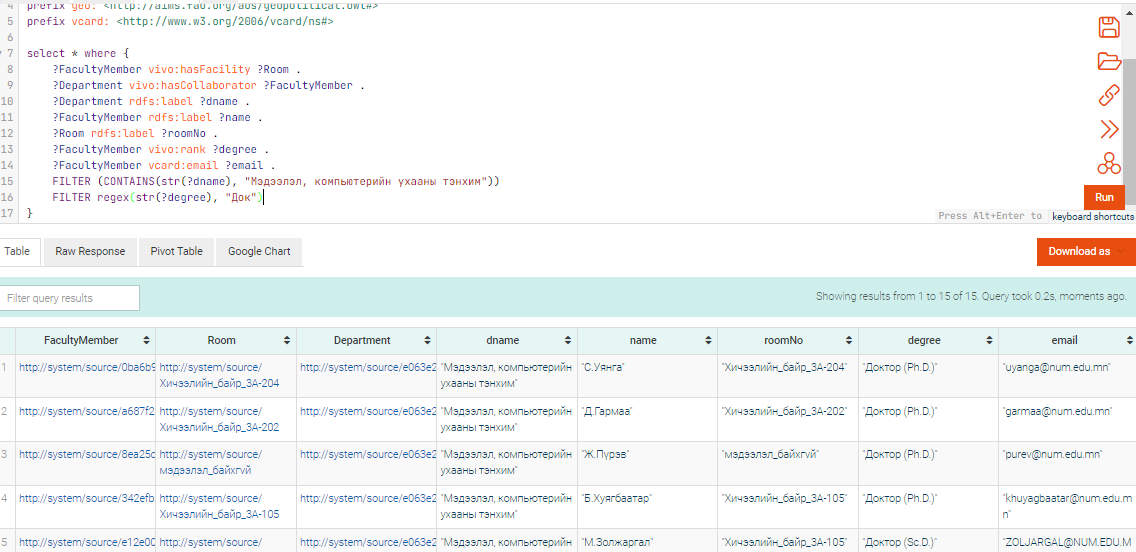
4.2.2 Мэдээлэл компьютерийн ухааны тэнхимийн бакалаврын хөтөлбөрүүдийг батлагдсан онтой нь харуулна уу?





4.2.3 Мэдээлэл компьютерийн ухааны тэнхимийн докторын зэрэгтэй багш, ажилчдын өрөө, email, нэрийг харуулна уу?





# 5. ДҮГНЭЛТ

Энэхүү лабораторын ажлыг хийснээр МУИС-ийн өгөгдлийн сангаас хүснэгтэн мэдээллийг татаж авах, Karma Data Integration tool, МУИС-ийн онтологийг ашиглан RDF график мэдээллийн сан үүсгэх, мэдээллийн санг GraphDB-тэй нэгтгэх зэрэг ажлуудыг хийсэн. Холбогдох мэдээллийг олж авахын тулд график мэдээллийн санд асуулга явуулсан. Даалгаврын үр дүн нь өгөгдлийг нэгтгэх, асуулга хийхэд семантик технологийг ашиглах талаархи ойлголтыг өгдөг. Хүснэгтийн өгөгдлийг RDF-ийн график мэдээллийн санд нэгтгэснээр МУИС-ийн мэдээллийн портал дахь мэдээллийн нөөцийг судлах, ашиглахад илүү үр дүнтэй мэдээлэл хадгалах, хайх, дүн шинжилгээ хийх боломжийг олгодог.

# 6. ХАВСРАЛТ

*prefix onto:<http://www.ontotext.com/>*

*prefix obo: <http://purl.obolibrary.org/obo/>*

*prefix vivo:<http://vivoweb.org/ontology/core#>*

*prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>*

*PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>*

*select \* where {*

*?FacultyMember obo:RO\_0000056 ?Course .*

*?FacultyMember vivo:hasFacility ?Room .*

*?FacultyMember rdfs:label ?lecturer .*

*?Room rdfs:label ?roomNo .*

*?Course rdfs:label ?course .*

*FILTER regex(str(?lecturer), "Хуягбаатар")*

*}*

*prefix obo: <http://purl.obolibrary.org/obo/>*

*prefix vivo:<http://vivoweb.org/ontology/core#>*

*prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>*

*prefix geo: <http://aims.fao.org/aos/geopolitical.owl#>*

*select \* where {*

*?department obo:ERO\_0000390 ?program .*

*?department rdfs:label ?name .*

*?program rdfs:label ?pname .*

*?program geo:validSince ?date .*

*?program vivo:rank ?rank .*

*FILTER (CONTAINS(str(?name), "Мэдээлэл, компьютерийн ухааны тэнхим"))*

*FILTER regex(str(?rank), "Бак")*

*}*

*prefix obo: <http://purl.obolibrary.org/obo/>*

*prefix vivo:<http://vivoweb.org/ontology/core#>*

*prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>*

*prefix geo: <http://aims.fao.org/aos/geopolitical.owl#>*

*prefix vcard: <http://www.w3.org/2006/vcard/ns#>*

*select \* where {*

*?FacultyMember vivo:hasFacility ?Room .*

*?Department vivo:hasCollaborator ?FacultyMember .*

*?Department rdfs:label ?dname .*

*?FacultyMember rdfs:label ?name .*

*?Room rdfs:label ?roomNo .*

*?FacultyMember vivo:rank ?degree .*

*?FacultyMember vcard:email ?email .*

*FILTER (CONTAINS(str(?dname), "Мэдээлэл, компьютерийн ухааны тэнхим"))*

*FILTER regex(str(?degree), "Док")*

*}*