

# Data Integration && SPARQL

У. Төрболд

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

## 1. 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

## 2. ЗОРИЛГО

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

## 3. ОНОЛЫН СУДАЛГАА

- 3.1. 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.
- 3.2. 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.
- 3.3. 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

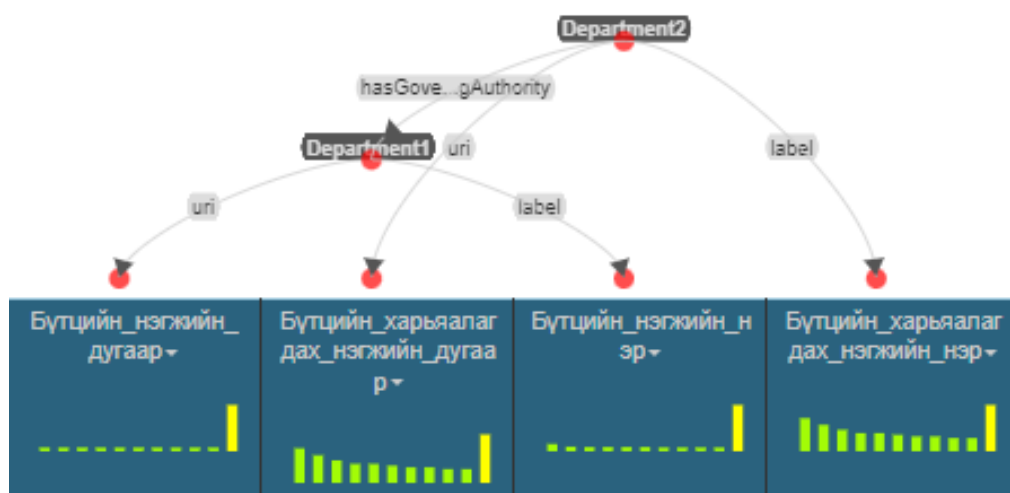
чанаруудын урьдчилан тодорхойлсон багцыг өгдөг. Хүснэгтийн өгөгдлийг онтологитой уялдуулах нь өгөгдлийн ойлголт, нэгтгэлийг сайжруулж, өөр өөр байгууллагуудын хооронд утга учиртай холболт, харилцааг бий болгоно.

3.4. 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.

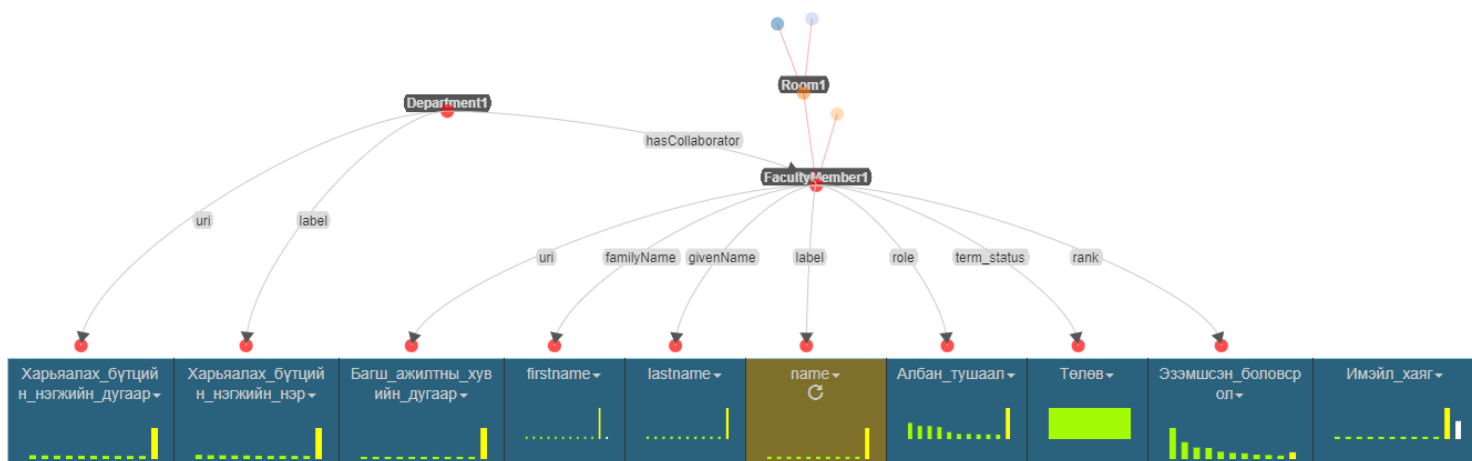
3.5. Select query: GraphDB ашиглан RDF график мэдээллийн санд дор хаяж гурван Select query боловсруулж, гүйцэтгэх нь график мэдээллийн санг нэгдсэн өгөгдлөөс утга учиртай, гүнзгий мэдээлэл олж авах чадварыг харуулах ёстой.

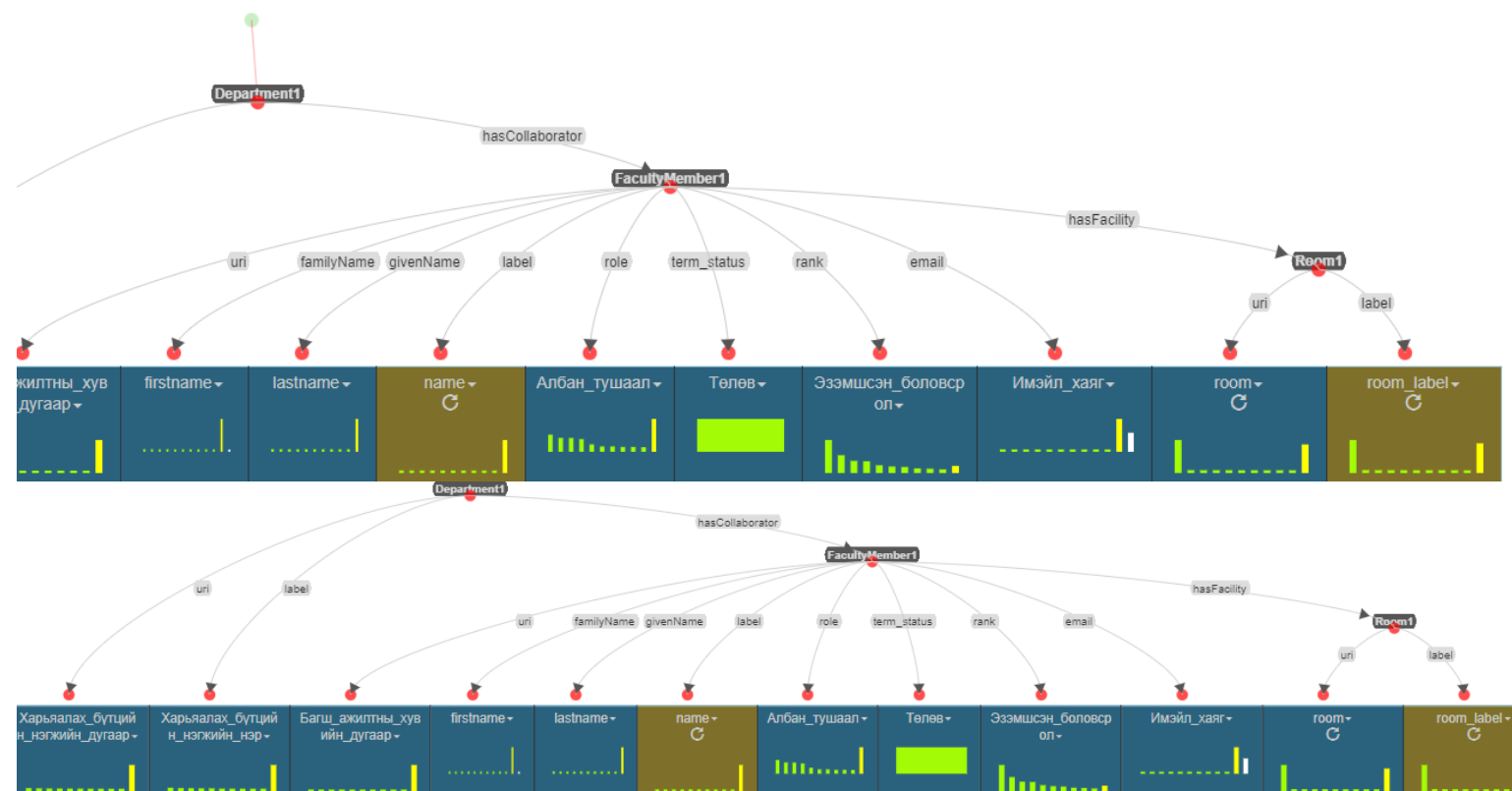
## 4. ХЭРЭГЖҮҮЛЭЛТ

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

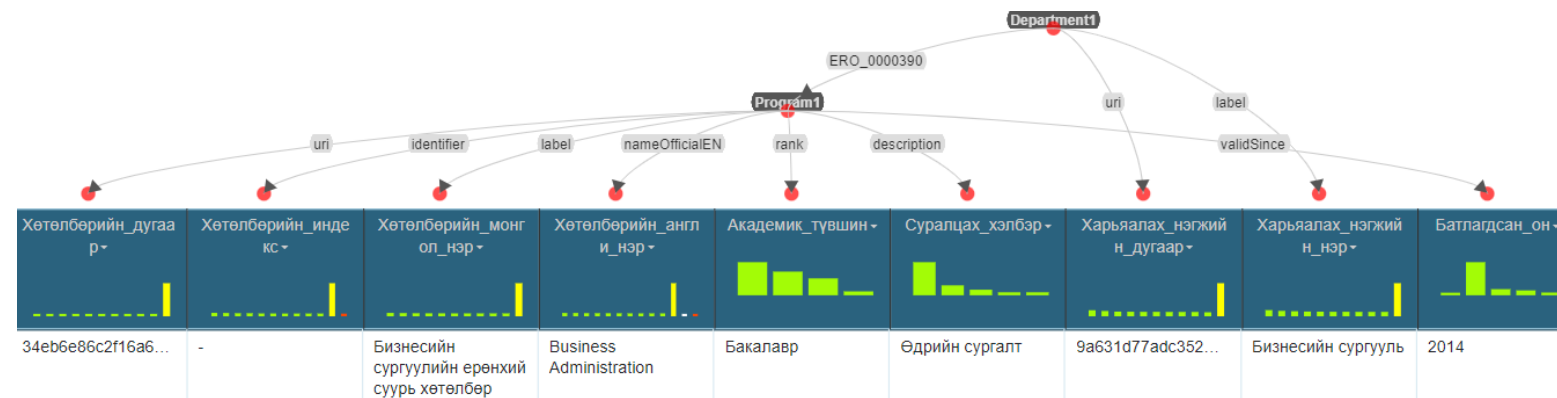


4.1.2. Багш ажилчдын нэрс өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал

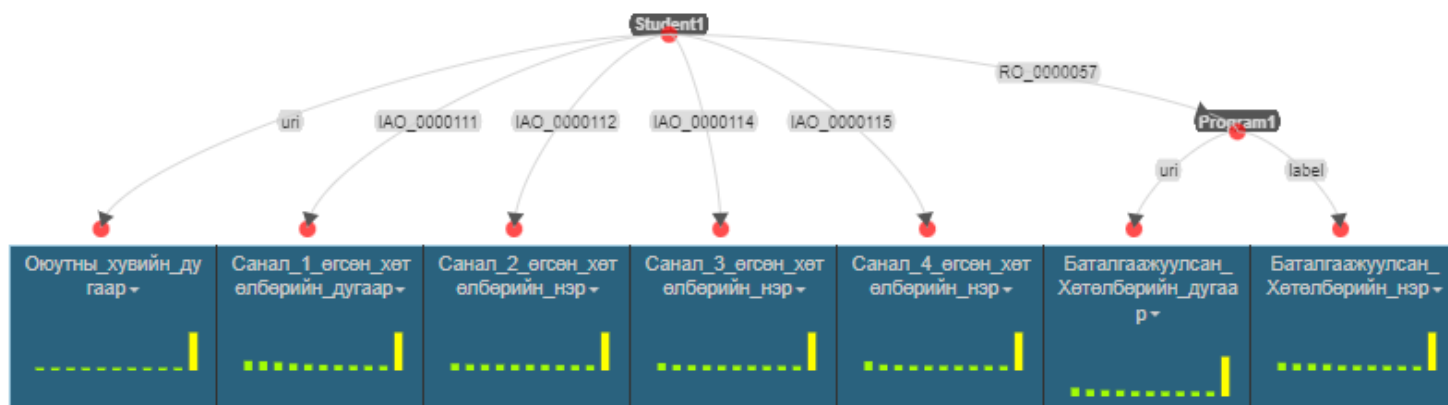




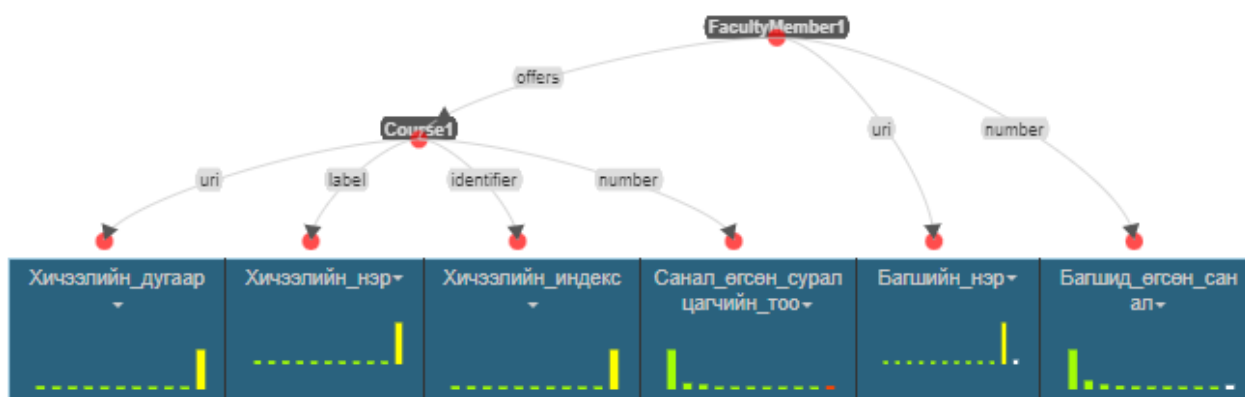
4.1.3. Сургалтын хөтөлбөр өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



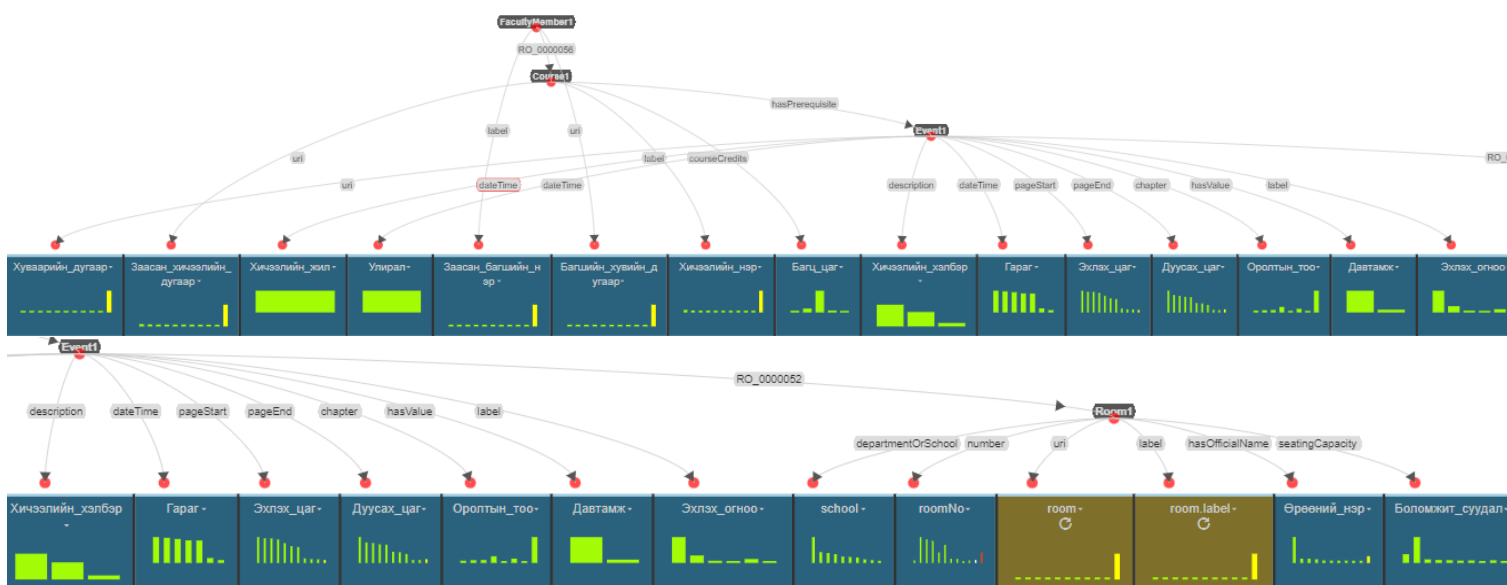
4.1.4. Хөтөлбөр сонголтын баталгаажуулалт өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал

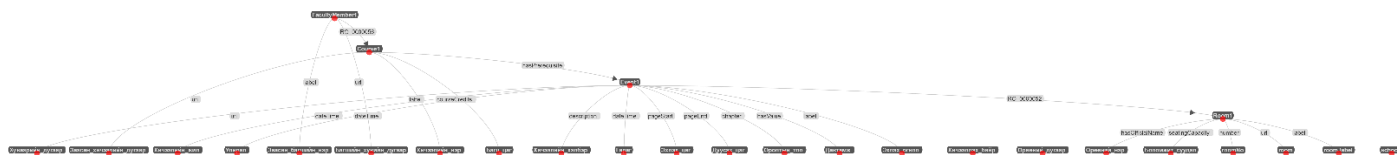


4.1.5. Хөтөлбөр сонгох санал өгөгдлийг Karma Data Integration tool болон vivo онтологи ашиглан RDF граф өгөгдлийн санг үүсгэсэн байдал



4.1.6. 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?

```

1 prefix onto:<http://www.ontotext.com/>
2 prefix obo: <http://purl.obolibrary.org/obo/>
3 prefix vivo:<http://vivoweb.org/ontology/core#>
4 prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
5 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
6
7 select * where {
8   ?FacultyMember obo:R0_0000056 ?Course .
9   ?FacultyMember vivo:hasFacility ?Room .
10  ?FacultyMember rdfs:label ?lecturer .
11  ?Room rdfs:label ?roomNo .
12  ?Course rdfs:label ?course .
13  FILTER regex(str(?lecturer), "Хуягбаатар")
14 }

```

```

1 prefix onto:<http://www.ontotext.com/>
2 prefix obo: <http://purl.obolibrary.org/obo/>
3 prefix vivo:<http://vivoweb.org/ontology/core#>
4 prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
5 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
6
7 select * where {
8   ?FacultyMember obo:R0_0000056 ?Course .
9   ?FacultyMember vivo:hasFacility ?Room .
10  ?FacultyMember rdfs:label ?lecturer .
11  ?Room rdfs:label ?roomNo .
12  ?Course rdfs:label ?course .
13  FILTER regex(str(?lecturer), "Хуягбаатар")
14 }

```

Run

keyboard shortcuts

Table

Raw Response

Pivot Table

Google Chart

Download as

Filter query results

Showing results from 1 to 9 of 9. Query took 2s. minutes ago.

	FacultyMember	Course	Room	lecturer	roomNo	course
1	http://system/source/342efb82a7bb2f	http://system/source/217b914abfc703	http://system/source/Хичээлийн_байр_3А-105	"Б.Хуягбаатар"	"Хичээлийн_байр_3А-105"	"Өгөгдлийн сангийн үндэс"
2	http://system/source/342efb82a7bb2f	http://system/source/f5da0f17e129a33	http://system/source/Хичээлийн_байр_3А-105	"Б.Хуягбаатар"	"Хичээлийн_байр_3А-105"	"Объект хандлагат програмчлал"
3	http://system/source/342efb82a7bb2f	http://system/source/996a0410bfc0f2	http://system/source/Хичээлийн_байр_3А-105	"Б.Хуягбаатар"	"Хичээлийн_байр_3А-105"	"Лекц"
4	http://system/source/342efb82a7bb2f	http://system/source/1f6533c21dd0ae	http://system/source/Хичээлийн_байр_3А-105	"Б.Хуягбаатар"	"Хичээлийн_байр_3А-105"	"Лаборатори"
5	http://system/source/342efb82a7bb2f	http://system/source/7ef4b99c31a7eb	http://system/source/...	"Б.Хуягбаатар"	"Хичээлийн_байр_3А-105"	"Лаборатори"

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

```

1 prefix obo: <http://purl.obolibrary.org/obo/>
2 prefix vivo:<http://vivoweb.org/ontology/core#>
3 prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 prefix geo: <http://aims.fao.org/aos/geopolitical.owl#>
5
6 select * where {
7     ?department obo:ERO_0000390 ?program .
8     ?department rdfs:label ?name .
9     ?program rdfs:label ?pname .
10    ?program geo:validSince ?date .
11    ?program vivo:rank ?rank .
12    FILTER (CONTAINS(str(?name), "Мэдээлэл, компьютерийн ухааны тэнхим"))
13    FILTER regex(str(?rank), "Бак")
14 }

```

Table Raw Response Pivot Table Google Chart

Filter query results Showing results from 1 to 4 of 4. Query took 0.2s

	department	program	name	pname	date	rank
1	http://system/source/e063e25fa020271	http://system/source/5eb5499d2e337e	"Мэдээлэл, компьютерийн ухааны тэнхим"	"Компьютерийн ухаан"	"2015"	"Бакалавр"
2	http://system/source/e063e25fa020271	http://system/source/2dfb5c5404b979e	"Мэдээлэл, компьютерийн ухааны тэнхим"	"Мэдээллийн систем"	"2015"	"Бакалавр"
3	http://system/source/e063e25fa020271	http://system/source/2492e5eaad61ec	"Мэдээлэл, компьютерийн ухааны тэнхим"	"Мэдээллийн технологи"	"2015"	"Бакалавр"
4	http://system/source/e063e25fa020271	http://system/source/d06e520797271a	"Мэдээлэл, компьютерийн ухааны тэнхим"	"Програм хангамж"	"2015"	"Бакалавр"

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

```

4 prefix geo: <http://aims.fao.org/aos/geopolitical.owl#>
5 prefix vcard: <http://www.w3.org/2006/vcard/ns#>
6
7 select * where {
8     ?FacultyMember vivo:hasFacility ?Room .
9     ?Department vivo:hasCollaborator ?FacultyMember .
10    ?Department rdfs:label ?dname .
11    ?FacultyMember rdfs:label ?name .
12    ?Room rdfs:label ?roomNo .
13    ?FacultyMember vivo:rank ?degree .
14    ?FacultyMember vcard:email ?email .
15    FILTER (CONTAINS(str(?dname), "Мэдээлэл, компьютерийн ухааны тэнхим"))
16    FILTER regex(str(?degree), "Док")
17 }

```

4	prefix vivo: <http://aims.tao.org/aos/geo/ontology#>
5	prefix vcard: <http://www.w3.org/2006/vcard/ns#>
6	
7	select * where {
8	?FacultyMember vivo:hasFacility ?Room .
9	?Department vivo:hasCollaborator ?FacultyMember .
10	?Department rdfs:label ?dname .
11	?FacultyMember rdfs:label ?name .
12	?Room rdfs:label ?roomNo .
13	?FacultyMember vivo:rank ?degree .
14	?FacultyMember vcard:email ?email .
15	FILTER (CONTAINS(str(?dname), "Мэдээлэл, компьютерийн ухааны тэнхим"))
16	FILTER regex(str(?degree), "Док")
17	}

Press Alt+Enter to keyboard shortcuts

Run

Table

Raw Response

Pivot Table

Google Chart

Download as

Filter query results

Showing results from 1 to 15 of 15. Query took 0.2s, moments ago.

	FacultyMember	Room	Department	dname	name	roomNo	degree	email
1	http://system/source/0ba6b59	http://system/source/Хичээлийн_байр_3А-204	http://system/source/e063e2	Мэдээлэл, компьютерийн ухааны тэнхим	О.Уянга	Хичээлийн_байр_3А-204	Доктор (Ph.D.)	uyanga@num.edu.mn
2	http://system/source/a687f2	http://system/source/Хичээлийн_байр_3А-202	http://system/source/e063e2	Мэдээлэл, компьютерийн ухааны тэнхим	Д.Гармаа	Хичээлийн_байр_3А-202	Доктор (Ph.D.)	garmaa@num.edu.mn
3	http://system/source/8ea25c	http://system/source/мэдээлэл_байр_гүй	http://system/source/e063e2	Мэдээлэл, компьютерийн ухааны тэнхим	Ж.Пүрэв	мэдээлэл_байр_гүй	Доктор (Ph.D.)	purev@num.edu.mn
4	http://system/source/342efb	http://system/source/Хичээлийн_байр_3А-105	http://system/source/e063e2	Мэдээлэл, компьютерийн ухааны тэнхим	Б.Хуягбаатар	Хичээлийн_байр_3А-105	Доктор (Ph.D.)	khuyagbaatar@num.edu.mn
5	http://system/source/e12e0c	http://system/source/	http://system/source/e063e2	Мэдээлэл, компьютерийн	М.Золжаргал	Хичээлийн_байр_3А-105	Доктор (Sc.D.)	ZOLJARGAL@NUM.EDU.MN

## 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), "Хүязбаатар")
}

```

```

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")
}
```