

МОНГОЛ УЛСЫН ИХ СУРГУУЛЬ
ХЭРЭГЛЭЭНИЙ ШИНЖЛЭХ УХААН, ИНЖЕНЕРЧЛЭЛИЙН СУРГУУЛЬ
МЭДЭЭЛЭЛ, КОМПЬЮТЕРИЙН УХААНЫ ТЭНХИМ

Дагдандоржийн Баярцэцэг

Монгол хэл дээрх чатбот
(Conversational chatbot for Multi-tenant)

Мэдээллийн систем (D061303)
Бакалаврын судалгааны ажил

Улаанбаатар

2022 он

МОНГОЛ УЛСЫН ИХ СУРГУУЛЬ
ХЭРЭГЛЭЭНИЙ ШИНЖЛЭХ УХААН, ИНЖЕНЕРЧЛЭЛИЙН СУРГУУЛЬ
МЭДЭЭЛЭЛ, КОМПЬЮТЕРИЙН УХААНЫ ТЭНХИМ

Монгол хэл дээрх чатбот
(Conversational chatbot for Multi-tenant)

Мэдээллийн систем (D061303)
Бакалаврын судалгааны ажил

Удирдагч: _____ Ph.D. М.Золжаргал

Гүйцэтгэсэн: _____ Д.Баярцэцэг (18B1NUM1306)

Улаанбаатар

2022 он

Зохиогчийн баталгаа

Миний бие Дагдандоржийн Баярцэцэг ”Монгол хэл дээрх чатбот” сэдэвтэй судалгааны ажлыг гүйцэтгэсэн болохыг зарлаж дараах зүйлсийг баталж байна:

- Ажил нь бүхэлдээ эсвэл ихэнхдээ Монгол Улсын Их Сургуулийн зэрэг горилохоор дэвшүүлсэн болно.
- Энэ ажлын аль нэг хэсгийг эсвэл бүхлээр нь ямар нэг их, дээд сургуулийн зэрэг горилохоор оруулж байгаагүй.
- Бусдын хийсэн ажлаас хуулбарлаагүй, ашигласан бол ишлэл, зүүлт хийсэн.
- Ажлыг би өөрөө (хамтарч) хийсэн ба миний хийсэн ажил, үзүүлсэн дэмжлэгийг дипломын ажилд тодорхой тусгасан.
- Ажилд тусалсан бүх эх сурвалжид талархаж байна.

Гарын үсэг: _____

Огноо: _____

МОНГОЛ УЛСЫН ИХ СУРГУУЛЬ
ХЭРЭГЛЭЭНИЙ ШИНЖЛЭХ УХААН, ИНЖЕНЕРЧЛЭЛИЙН СУРГУУЛЬ
МЭДЭЭЛЭЛ, КОМПЬЮТЕРИЙН УХААНЫ ТЭНХИМ

Монгол хэл дээрх чатбот
(Conversational chatbot for Multi-tenant)

Мэдээллийн систем (D061303)
Бакалаврын судалгааны ажил

Удирдагч:

М. Золжаргал

Ph.D. М.Золжаргал

Гүйцэтгэсэн:

Д. Баярцэцэг

Д.Баярцэцэг (18B1NUM1306)

Улаанбаатар

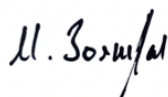
2022 он

**МУИС, ХШУИС-ийн Мэдээллийн Систем хөтөлбөрийн оюутан Д.Баярцэцэгийн
бакалаврын судалгааны ажлын удирдагчийн тодорхойлолт**

Оюутан Баярцэцэг нь нэгдсэн нэг цонхноос олон хувь хүн эсвэл компанид зориулсан чатбот үүсгэх, бүртгэх программ хангамжийн үйлчилгээг бакалаврын судалгааны ажлын хүрээнд амжилттай гүйцэтгэсэн бөгөөд чат бот болон хэл боловсруулалтын талаар суралцан санаачлагатай ажилласан. Ажлын хүрээнд машин сургалтад суурилсан чатботыг үүсгэсэн.

Оюутан нь судалгааны ажлын зорилгодоо хүрэхийн тулд хэл боловсруулалт болон javascript суурьтай веб хөгжүүлэлтийн технологийг бие даан судалж идэвх чармайлттай ажилласан учир удирдагчийн зүгээс 95% (А) үнэлгээ өгөх саналтай байна.

Удирдагч:



Д-р. М.Золжаргал

МУИС, ХШУИС, МКУТ-ийн дэд проф.

2022 оны 12 сарын 12

Мэдээллийн системийн хөтөлбөрийн оюутан
Дагдандоржийн овогтой Баярцэцэгийн “Монгол хэл дээрх чатбот” сэдэвт
бакалаврын дипломын ажлын шүүмж

Оюутан Д. Баярцэцэг бакалаврын дипломын ажлаараа Монгол хэлээр, ухаалаг, хялбар туршилтын чатбот хөгжүүлж фэйсбүүк интерфэйстэй холбох боломжтой мөн олон хэрэглэгч чатбот үүсгэх боломжтой архитектур гаргаж турших зорилготой ажил байна.


Дипломын ажлын 5 бүлэг, ном зүй болон кодын хавсралт ороод нийт 29 нүүр хуудас тайлан бичсэн байна. Дипломын ажлын тайлангийн бүтэц 1-р бүлэг сэдэвтэй холбоотой судалгаа, 2-р бүлэг системийн шинжилгээ, 3-р бүлэгт системийн зохиомж хэсэг байсан бол 4-р бүлэгт хэрэгжүүлэлт, үр дүнг танилцуулсан бөгөөд дүгнэлт, ном зүй хавсралт гээд 5 дугаар бүлгийг кодын хэрэгжүүлэлтээ нэрэлсэн байгаа нь гарчиглал хийхдээ алдаа гаргасан байна гэж бодлоо.

Дипломын ажилтай бүхэлд нь танилцахад дипломын үндэслэл, зорилт дээр тодорхой ойлгомжтой бичсэн байна. Нэгдүгээр бүлэгт судалгаа хэтэрхий товч богино гэж үзэж болохоор бол ижил төстэй системтэй харьцуулахдаа тодорхой үзүүлэлтүүдээр харьцуулаагүй мөн зөвхөн зураг харуулсан хангалтгүй харьцуулалт хийсэн байна. Бусдаар хийх шаардлагатай систем шаардлага тодорхойлох, архитектур зохиомж хөгжүүлэлтийн ажлуудыг зохих түвшинд хийсэн байна. Системийн хэрэгжүүлэлтийн хувьд хялбархан хийсэн ер нь бол туршилтын энгийн ажил болсон байна гэж харагдаж байна.

Шүүмжлэгчийн зүгээс оюутан Д. Баярцэцэгийн дипломын ажилд 7 гэсэн үнэлгээ өгч байна.

Шүүмж бичсэн

Хэрэглээний Шинжлэх Ухаан, Инженерчлэлийн
Сургуулийн Мэдээлэл, Компьютерийн Ухааны
Тэнхимийн багш

З.Цолмон /  /

ГАРЧИГ

УДИРТГАЛ.....	2
1. СУДАЛГАА	3
1.1 Онолын хэсэг	3
1.2 Ижил төстэй программын судалгаа	5
1.3 Технологийн судалгаа	6
1.4 Бүлгийн дүгнэлт	11
2. СИСТЕМИЙН ШИНЖИЛГЭЭ.....	12
2.1 Функционал ба функционал бус шаардлагууд	12
2.2 Динамик загвар	13
2.3 Бүлгийн дүгнэлт	15
3. СИСТЕМИЙН ЗОХИОМЖ	16
3.1 Системийн ерөнхий архитектур	16
4. ХЭРЭГЖҮҮЛЭЛТ, ҮР ДҮН.....	17
4.1 Rasa фреймворк	17
4.2 Чатботын хэрэгжүүлэлт	20
4.3 Бүлгийн дүгнэлт	22
ДҮГНЭЛТ	23
НОМ ЗҮЙ	23
ХАВСРАЛТ	24
А. КОДЫН ХЭРЭГЖҮҮЛЭЛТ	25
А.1 Action хэсгийн код	25
А.2 Харилцан ярианы түүх хэсэг	26
А.3 Хайх lookur жагсаалт	28
А.4 Ботын хариулах хэсгийн код	29
А.5 Ботын дүрэм боловсруулалт	31

ЗУРГИЙН ЖАГСААЛТ

1.1	Multi-tenant төрөл	3
1.2	Multi-tenant архитектур	4
1.3	Manychat нь дараах программтай хамтарч ажилладаг.	5
1.4	Ижил төстэй технологийн харьцуулалт	6
1.5	Docker-н ажиллагаа	8
2.1	Чатботын асуулт хариултын үйл явц	13
2.2	Чатботын үйл ажиллагааны диаграмм	15
3.1	Чатботын ерөнхий архитектур	16
3.2	Чатботын архитектур	16
4.1	Rasa үндсэн үйл ажиллагаа	17
4.2	Зохиол ба дүрмийн жишээ	19
4.3	Чатбот хэрэглэгчийн endpoint	20
4.4	Туршилтын энгийн ботын үр дүн	20
4.5	Туршилтын динамик шинжтэй ботын үр дүн	20
4.6	Туршилтын ботын үр дүн	21

Товчилсон үгийн жагсаалт

Short message service (SMS) - Богино мессежийн үйлчилгээ

Natural Language Processing (NLP) - Хэлний боловсруулалт

Natural Language Understanding (NLU) -Эх хэлний ойлголт

Natural Language Generation (NLG) - Эх хэлний үүсгэлт

Artificial Intelligence (AI) - Хиймэл оюун ухаан

УДИРТГАЛ

Байгууллагуудын үйл ажиллагааг автоматжуулж буй нэгэн хэрэгсэл бол харилцан ярианы интерфейс - чатбот болжээ. Үүний ачаар бизнесийн үйлчилгээ болон үйл ажиллагаа нь хэрэглэгчдэд төвөггүй, ойлгомжтой, 24/7 цагаар хүрч чадах юм. Дэлхийн улс орнууд үүнийг аль хэдийн ашиглаад эхэлсэн байна.

Ихэнхи тохиолдолд хэн бүхний хэрэгцээт мэдээлэлээ олох үйл явц нь маш төвөгтэй, удаан байдаг. Аливаа мэдээллийн тухайн гарын авлага, нийтлэл, социал хаяг, веб хуудас гэх мэт сувгууд түгээмэл олон байдаг ч богино хугацаанд, үнэн зөв, хүссэн мэдээлэлээ олно гэдэг эргэлзээтэй.

Тэгвэл энэ асуудлыг чатбот хялбархан шийдэж чадна. Миний бие хиймэл оюун ухаанд суурилсан хүний хэлийг ойлгодог олон төрлийн мэдээллийг нэг дороос хүргэх чадвартай чатботын хөгжүүлэлт, түүний өргөтгөх боломжийн талаар судалж, судалсан мэдлэгээ бататган турших юм.

Зорилго

Энэхүү өөрийн бакалаврын судалгаат ажилын хүрээнд Монгол хэл дээрх ухаалаг, хялбар демо чатбот хөгжүүлж фэйсбүүк интерфэйстэй холбох ба олон хэрэглэгч чатбот үүсгэх боломжтой архитектур гаргаж, олсон шийдлээ туршин хэрэгжүүлэхэд оршино. Зорилгодоо үндэслэн дараах зорилтуудыг дэвшүүлэв:

- хэрэглэгчийн эцсийн төгсгөл(endpoint) үүсгэх
- DialogFlow судлан хэл боловсруулалтын талаар ойлголттой болох
- Rasa нээлттэй эхийн платформыг судлах
- Демо чатботын хэрэгжүүлэлт хийх
- Демогоо интерфэйстэй холбох

Мөн төлөвлөгөөний дагуу хөгжүүлэлтэнд явагдах ажлуудыг гүйцэтгэх юм.

1. СУДАЛГАА

1.1 Онолын хэсэг

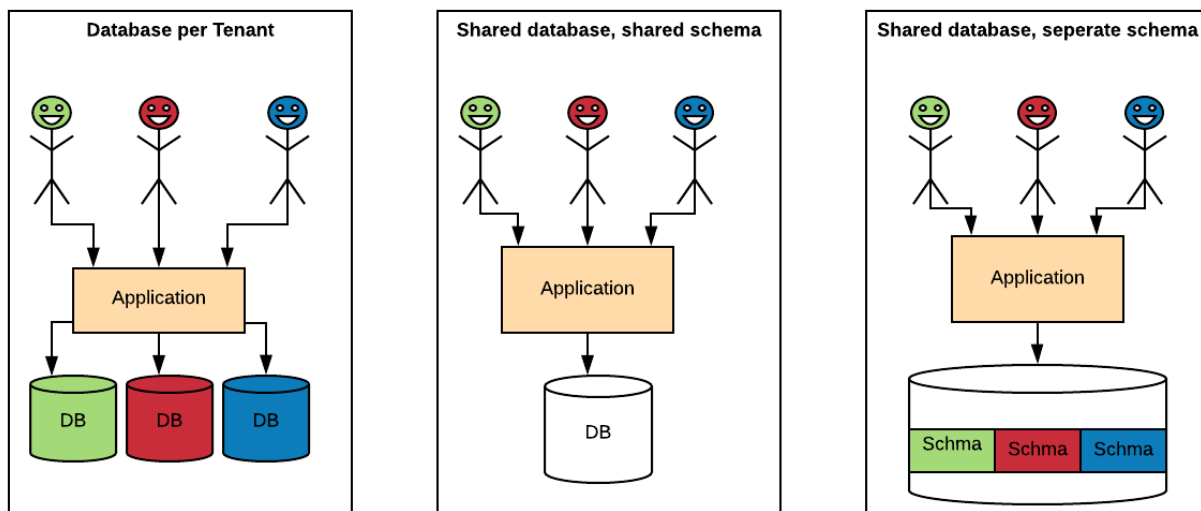
Энэхүү бүлэгт чатбот, multi-tenant архитектур, түүний төрөл, онцлогын тухай авч үзнэ.

1.1.1 Multi-tenant

Multi + tenant

Түрээслэгч (tenant) гэдэг нь нэг байгууллага, компанийн хэрэглэгчидийн бүлэг гэж тодорхойлж болно. Олон хэрэглэгчид (multi-tenant) тус бүр нэг программ хангамж болон түүний дэд бүтэц мэдээллийн санг хуваалцахыг хэлнэ. Multi-tenant нь дараах гурван төрөлтэй.

Зураг 1.1: Multi-tenant төрөл



Дараах архитектураар хэрэгжүүлсэн программын жишээг дор жагсаав. Жишээ нь: Gmail, Github г.м программууд бий.

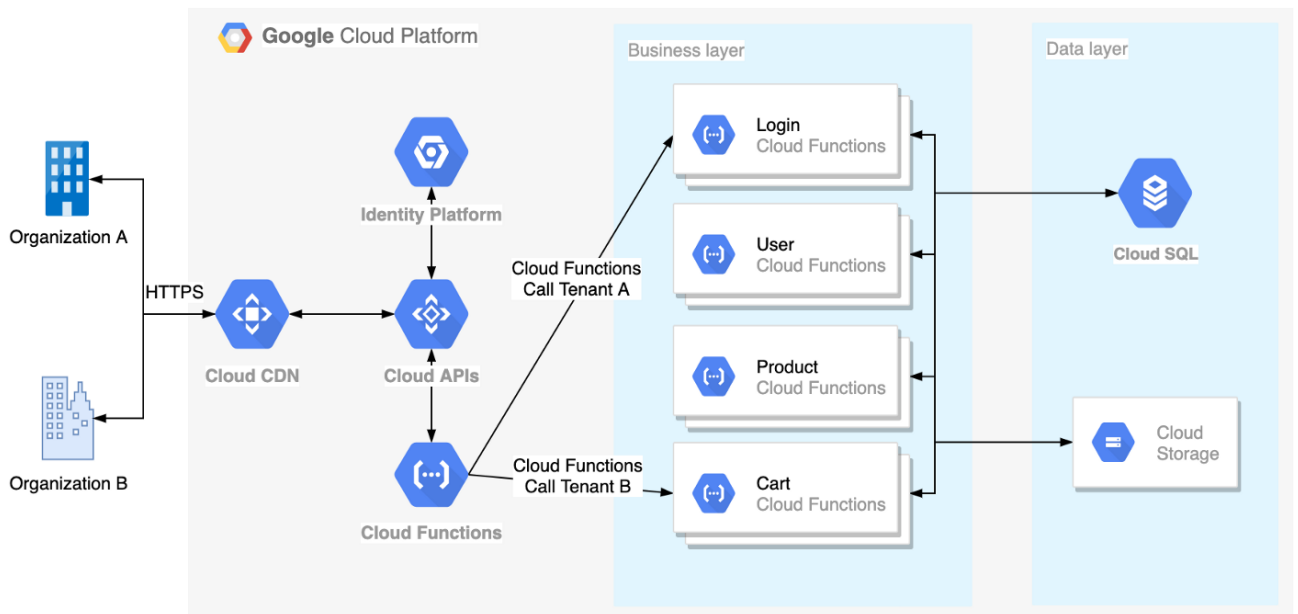
Multi-tenant давуу талыг жагсаавал:

- Архитектурыг хялбархан өргөжүүлэх боломжтой.
- Үйлчилгээ үзүүлэгчид зөвхөн нэг системийг хянаж, удирдах ёстой.

Multi-tenant сул талыг жагсаавал:

- Олон түрээслэгчтэй программууд аюулгүй байдлын үүднээс илүү хатуу баталгаажуулалт, хандалтын хяналт шаарддаг.
- Олон түрээслэгчтэй программууд нь бусад түрээслэгчийн архитектурын программуудаас бага уян хатан байх хандлагатай байдаг.
- Сүлжээний тасалдал үүсэж болно.
- Олон хэрэглэгч дахь тухайн нэг хэрэглэгч нь хэт их ачааллуулж байвал вэб хостинг , мэдээллийн сан , сүлжээ , хадгалалт , серверт нөлөөлж сүлжээний гүйцэтгэлийн асуудал үүсгэж болно.[1]

Зураг 1.2: Multi-tenant архитектур



1.1.2 Чатбот

Чатбот бол хүний яриаг дуурайлган хийх зориулалттай программ хангамж юм. Олон чатботууд мессежийн програмууд эсвэл богино мессежийн үйлчилгээ (SMS) дээр ажилладаг. Чатбот технологи нь хүн төрөлхтөнд юу хэрэгтэйг ойлгож, эцсийн хэрэглэгчийг хүссэн үр дүнд хүргэхэд туслах хариу үйлдлийг үзүүлэх юм. Түүндээ хэлний боловсруулалт (NLP) болон хиймэл оюун ухааныг (AI) ашигладаг.

Чатботын онцлог:

- Маш бага зардлаар 24/7 ажиллана.
- Хэрэглэгчдэд хэрэгтэй, үнэн бодит мэдээлэл өгнө.
- Захиалга авах, төлбөр тооцоо хийх хүртэл харилцагчийн үйлчилгээг үзүүлнэ.
- Хурдан шуурхай, тогтмол асуултанд хариулдаг.
- Хүнээс хамаарсан алдаа гарах нь багасна.
- Хэрэглэгчийн сэтгэл ханамж ихэснэ.

1.2 Ижил төстэй программын судалгаа

1.2.1 ManyChat

Manychat бол хамгийн их хэрэглэгддэг чат маркетингийн платформ юм. Код бичих шаардлагагүй - гээр харилцан яриаг хялбархан бүтээж, автоматжуулдаг. [2]

Зураг 1.3: Manychat нь дараах программтай хамтарч ажилладаг.



Messenger
Facebook Messenger
chatbot #1



Instagram new
Automate your
Instagram Marketing



SMS
Text with Customers



WhatsApp
Connect with your
Customers Instantly



Integrations
Connect to your
favorite tools

1.2.2 Chatfuel.com

Chatfuel нь Facebook, Instagram-д холбогдох боломжтой код бичих шаардлагагүй бот бүтээх платформ. Энэ технологи нь жижиг бизнесээс эхлээд Adidas, Levi's зэрэг томоохон брэндүүд хүртэл чатын автоматжуулалтаар дамжуулан үйлчилгээгээ явуулдаг.

1.2.3 Chatbot.mn

Монгол AI чатбот платформ. Энэхүү chatbot.mn-ийг ашиглан хэрэглэгчийн асуултыг ойлгодог, төлбөр хүлээн авдаг, хэрэглэгчийн зан төлөв дээр тулгуурлан суралцдаг чатботыг үнэ төлбөргүй хийх боломжтой. Хамгийн гол нь ямар нэг программчлалын чадвар шаардахгүй.

Зураг 1.4: Ижил төстэй технологийн харьцуулалт

	Chatbot.mn	Chatfuel	ManyChat
Хэрэглэгчийн хувьд	меню цэс	Хяналтын самбар анализ мэдээлэлээр баялаг	Ашиглахад хялбар
Элемэнт		Элемэнт: файл, босоо жагсаалт болон байршил хуваалцах боломжийг олгодоггүй.	Элемэнт: босоо жагсаалт оруулах боломжгүй. Tag ашиглах боломжгүй
AI		DialogFlow нь хиймэл оюун ухаан ашиглах боломжийг олгодог.	DialogFlow нь хиймэл оюун ухаан ашиглах боломжийг олгодог.
Хэл	Монгол хэл	Англи хэл	Англи хэл
Хамтрах платформ	Вэбсайт, Facebook messenger дээр чатбот үүсгэх боломжтой.	Вэбсайт, Facebook Messenger, Instagram дээр чатбот үүсгэх боломжтой.	Вэбсайт, Facebook Messenger, Instagram, Whatsapp, SMS, Email дээр чатбот үүсгэх боломжтой.

1.3 Технологийн судалгаа

Энэ бүлэгт програмын зорилгодоо үндэслэн ашиглаж болох хэрэглэгч ба сервер талыг хэрэгжүүлэх технологийг судлан харьцуулав.

1.3.1 Node js

Java script дээр яагаад сервер талын кодыг бичиж болохгүй гэж гэсэн асуултаас анх node js -ын үүсэл эхэлсэн бөгөөд энэ технологийг 2009 онд Ryan Dahl анх танилцуулсан ба үүнээс хойш хөгжсөөр байна. NodeJS нь ямар нэг фреймворк эсвэл номын сан биш харин Javascript кодыг бровсероос гадна ажиллуулж буй нээлттэй эхийн, олон платформыг дэмждэг ажиллах орчин юм.

Мөн NodeJS-ийн хэд хэдэн давуу талууд байна. Тухайлбал,

- Маш хурдан процесс хийдэг.

NodeJS нь хурдан байхын тулд non-blocking input/output болон asynchronous request handling ашигладаг. Энэ нь NodeJS-ийг ямар нэг хугацааны сааталгүйгээр хүсэлтүүдэд хариу өгөх, боловсруулах боломжтой болгодог. Сервер талд synchronous processing нь кодыг яг дарааллаар нь ажиллуулдаг. Ингэснээр хүсэлт бүрт зөвхөн нэг thread оноогддог ба бусад хүсэлтүүд нь дуустал нь хүлээх шаардлагатай болдог. Харин asynchronous processing нь нэг thread олон хүсэлт боловсруулахыг зөвшөөр- дөг.

- Microservice архитектурыг ашигладаг.

Нэг том monolithic системээс илүүтэй бизнес логикуудаа жижиг модулиудад хуваасан microservice архитектур нь цаашдыг өсөлт хөгжил болон уян хатан байдалд тустай.[3]

- Баялаг хөгжүүлэлтийн орчинтой.

NodeJS-ийн анхдагч package manager, npm нь технологийн хөгжлийн нэг чухал хэсэг болох нээлттэй эхийн Javascript хэрэгслүүдийг агуулдаг. Npm дээр одоогоор бүртгэлтэй байгаа хэрэглэх боломжтой 836000 сан байдаг бөгөөд долоо хоног бүр шинээр 10000 сангууд болон хэрэгслүүд нэмэгдэж байдаг. Энэ нь ямар баялаг экосистем, хөгжүүлэлтийн орчинтой болохыг харуулж байна.

- JSON дэмждэг.

JSON format-ийг binary model руу хөрвүүлэхгүйгээр ашигладаг. NoSQL өгөгдлийн сантай харьцдаг RESTful APIs-тай ажиллах шаардлагатай үед энэ ихээхэн тус болдог.

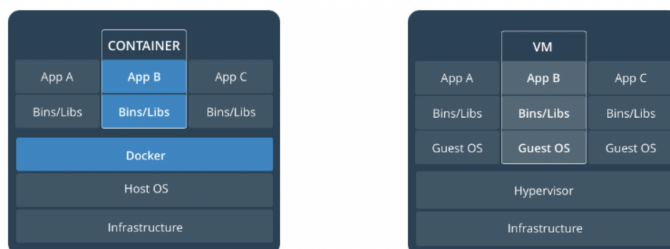
1.3.2 Docker

Docker бол програмыг хурдан бүтээх, турших, ашиглах боломжийг олгодог програм хангамжийн платформ юм. Docker нь програм хангамжийг номын сан, системийн хэрэгсэл, код ажиллах хугацаа гэх мэт программ хангамжийг ажиллуулахад шаардлагатай бүх зүйлийг агуулсан контейнер гэж нэрлэгддэг стандартчилсан нэгжүүдэд багцлагддаг. Docker-ийг ашигласнаар та ямар ч орчинд програмуудыг хурдан байрлуулж, масштабтай болж таны код ажиллах болно. [4]

Docker хэрхэн ажилладаг вэ

Docker бол контейнерт зориулсан үйлдлийн систем юм. Виртуал машин нь серверийн техник хангамжийг хэрхэн виртуалчилдаг (шууд удирдах хэрэгцээг арилгадаг)- тай адил контейнерууд нь серверийн үйлдлийн системийг виртуалчилдаг. Docker нь сервер бүр дээр суурилагдсан бөгөөд контейнер бүтээх, эхлүүлэх, зогсооход ашиглаж болох энгийн командуудыг өгдөг.

Зураг 1.5: Docker-н ажиллагаа



1.3.3 Mongo DB

MongoDB бол NoSQL технологид суурилсан нээлттэй эхийн өндөр хурдны мэдээллийн сан бөгөөд C++ хэл дээр хөгжүүлэгдсэн байдаг. Гүйцэтгэл, хүртээмж өндөр, хялбар цар

хүрээтэй үйлчилгээ үзүүлдэг.

Mongodb-ийн давуу талыг жагсаав. Тухайлбал:

- Уян хатан мэдээллийн сантай.

Mongodb нь схем-лэс өгөгдлийн сан юм. Document-ээс тусад нь ямар ч төрлийн өгөгдлийг хадгалж чаддаг. Энэ нь бидэнд янз бүрийн төрлийн өгөгдөл хадгалах уян хатан байдлыг өгдөг.

- Өндөр хурд

Mongodb нь document-oriented өгөгдлийн сан бөгөөд дугаарлалт буюу id-аар нь document руу хандахад амархан байдаг. Түүнчлэн маш хурдан query response-оор хангадаг. Mongodb-ийн хурд нь relational өгөгдлийн сангаас 100 дахин хурдан байдаг.

- Өргөтгөх боломжтой.

Mongodb-ийн бас нэг том давуу тал нь ачааллаас хамааран серверийн тоог нэмж хасах юм. Их хэмжээний өгөгдөл хадгалах хэрэгтэй тохиолдолд хэд хэдэн сервер дээр тарааж хадгалж болно.[5]

1.3.4 NLP

Хиймэл оюун ухааныг өөртөө хэрэгжүүлсэн чатботууд нь хэрэглэгчдэд илүү сайн туршлага өгөхийн тулд эх хэл боловсруулах (NLP) хэмээх хиймэл оюун ухааны салбарыг ашигладаг. Түүний ачаар чатбот илүү тохирсон, өвөрмөц хариултуудыг оновчтойгоор өгч, шинэ асуулт эсвэл тушаалуудыг ойлгож, хариулах, хэрэглэгчийн хэрэгцээнд нийцүүлэн хэрэглэгчийн туршлагыг сайжруулах боломжтой болдог. Эх хэлний боловсруулалт (NLP) гэдэг нь энгийн яриа болон текстээр программ хангамжийг автоматаар удирдах гэж ерөнхийд нь тодорхойлдог.

- Эх хэл нь хүмүүсийн бие биетэйгээ харилцахдаа хэрэглэдэг аргыг хэлдэг. Хүний хэл нь байнга өөрчлөгдөж, хөгжиж байдаг, уян хатан, төвөгтэй, нарийн мэдрэмжээр дүүрэн байдаг болохоор маш төвөгтэй юм. Чатбот хөгжүүлэгчид хүмүүстэй харилцах харилцааны

төлөөлөгчийг зохион бүтээхээсээ өмнө хүмүүс хоорондоо хэрхэн ярилцаж байгааг ойлгох нь маш чухал юм. Бидний тархи эх хэлээ бүтээх, шинэчлэх, ямар ч нарийн утгатай санааг ойлгох, илэрхийлэх, эргэцүүлэх, тайлбарлах чадвартай байдаг ч хэлний дагаж мөрддөг дүрмийг албан ёсоор хэн нэгэнд ойлгуулж, тайлбарлахдаа муу байдаг.

NLP нь хэл шинжлэл, компьютерын шинжлэх ухаан, хиймэл оюун ухаан гэх талбаруудын дэд салбар бөгөөд компьютер болон хүний хэл хоорондоо хэрхэн ойлголцох, өгөгдөл дээр бие даан ажиллах талаар судалдаг буюу гол төвлөрөл нь хүний илгээсэн мэдээллийг ”ойлгох” чадвартай компьютерыг бүтээх юм.

Хэрвээ команд нь ярианаас бүрдсэн тохиолдолд яриаг текст рүү хөрвүүлэх ажлыг гүйцэтгэдэг. NLP нь хоёр процессыг агуулдаг.

- **Эх хэлний ойлголт**

(Natural Language Understanding) - нь эх хэлийг ашиглан компьютертой харилцах боломжийг олгодог бөгөөд өгөгдсөн текстийн утгыг машинаар уншигдахуйц Формат болгон хувиргадаг хэсэг юм. Илтгэгчийн хэлсэн үгийн цаана байгаа сэтгэл хөдлөл, зорилго гэх мэт хүний харилцаанд байдаг олон утга санааг мэдрэхийг эрмэлздэг. Тэдгээрийг хэрэглэгчийн зорилго таних (intent extraction) болон нэгж- объектыг таних (entity extraction) гэнэ. Нэгж нь хэрэглэгчийн илгээсэн текстээс авах, харилцан ярианд нөлөөлөх мэдээллийг илтгэнэ. Жишээ нь, цаг агаар хэлдэг ботод хэрэглэгч аль хотын цаг агаарыг асууж байгаа нь чухал тул ”хот” нь нэгж юм. NLU текстийг машинд ойлгогдохуйц болгохын тулд тухайн хэлний хэв маяг, өвөрмөц хэллэгүүдийг болон текст доторх үг бүрийн мөн чанар, утга санаа, бүтцийг мэддэг байх ёстой. NLU нь ахисан түвшинд яриандаа үгний буруу дуудлага, бичигт солигдсон үсэг, буруу бичигдсэн үг зэрэг хүний нийтлэг алдаанаас ул хамааран утгыг ойлгохоор програмчлагдсан байдаг. Эдгээр машин сургалтын загварууд цаг хугацааны явцад сайжирдаг.

- **Эх хэлний үүсгэлт**

(Natural Language Generation) - өгөгдлийн багцаас эх хэлний гаралтыг гаргадаг хангамжийн процесс. Тухайн хэсэгт компьютер нь хүний хэлийг дуурайн уян хатан, сонирхол татахуйц текст бэлтгэх нь зорилго юм. Мэдээллээ бүрдүүлэх, өгүүлбэрийн бүтцээ тодорхойлох, үгээ сонгох, өгүүлбэрүүдээ нийлүүлээд хэрэглэгчрүү гаралт болгох зэрэг шатуудаас бүрдэнэ. олно.[6]

1.4 Бүлгийн дүгнэлт

Уг бүлгийг нэгтгэн дүгнэвэл дараах байдлаар тодорхойлно.

Хэрэглэгч тал: JavaScript + ReactJS + React Native.

Сервер тал: NodeJS + MongoDB + Rasa + Docker.

2. СИСТЕМИЙН ШИНЖИЛГЭЭ

2.1 Функционал ба функционал бус шаардлагууд

Тухайн програмыг хувь хүн, албан байгууллага ашиглах боломжтой. Доорх хэсэгт функционал ба функционал бус шаардлагуудыг жагсаав.

2.1.1 Функционал шаардлага

- ФШ1 Хэрэглэгч асуулт асууна
- ФШ2 Систем тохирох хариултыг өгдөг байна.
 - ФШ10 Хэрэглэгч бүртгэл үүсгэдэг байна.
 - ФШ12 Нэвтрэх эрхтэй байна.
 - ФШ13 Програмаас гардаг байна.

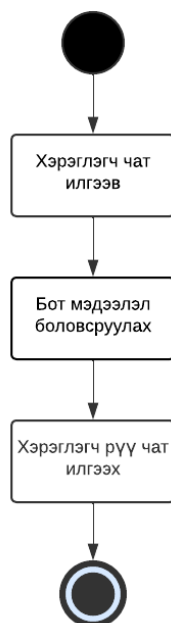
2.1.2 Функционал бус шаардлага

- ФБШ10 Явуулсан зурвас найдвартай илгээгддэг байна.
- ФБШ12 Интернет сүлжээнд холбогдож хэрэглэгддэг байна.

2.2 Динамик загвар

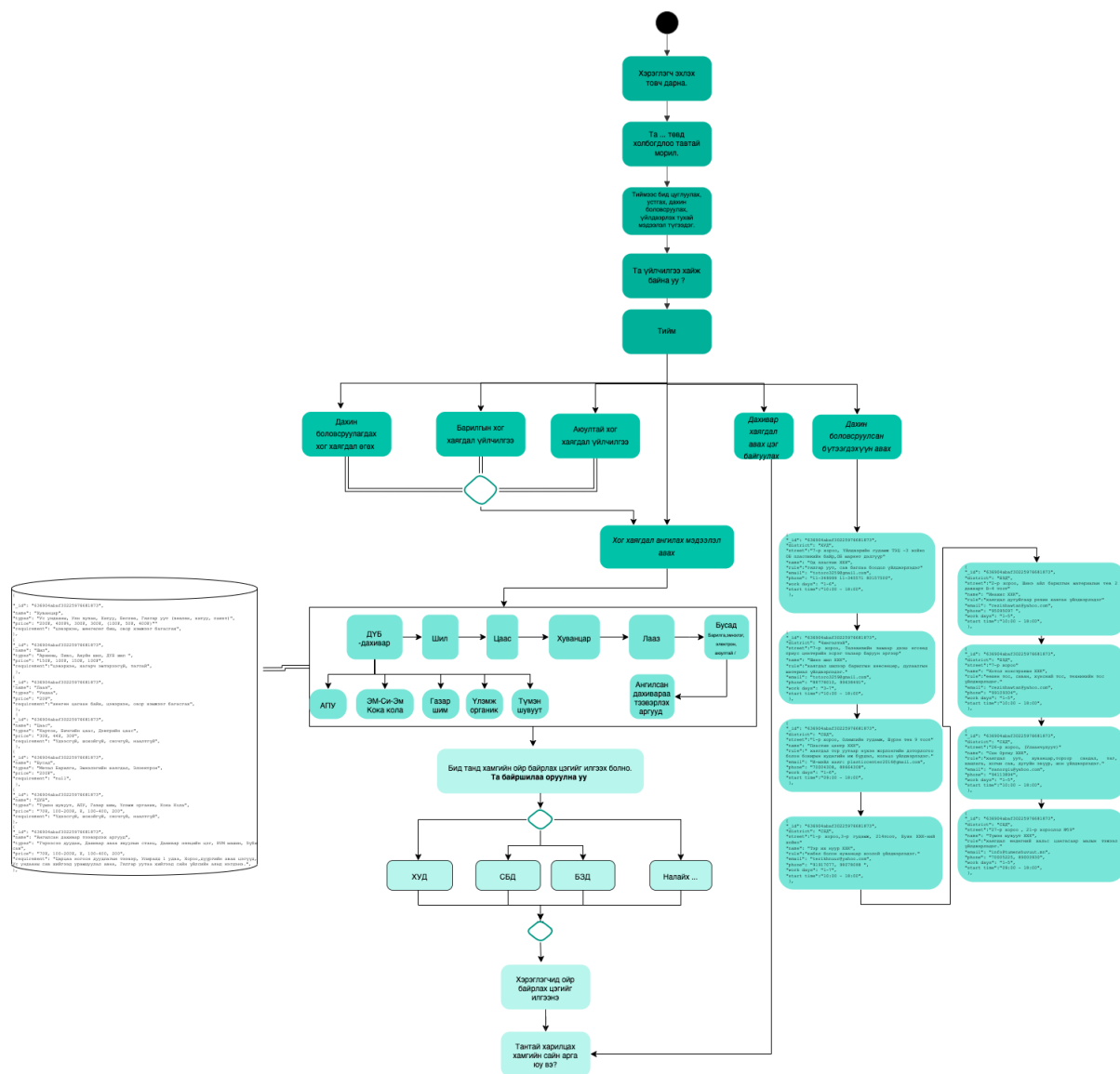
Програмын динамик загварыг ажлын явцын диаграмм болон үйл ажиллагааны диаграммаар дүрслэн харуулав.

2.2.1 Функционал загвар



Зураг 2.1: Чатботын асуулт хариултын үйл явц

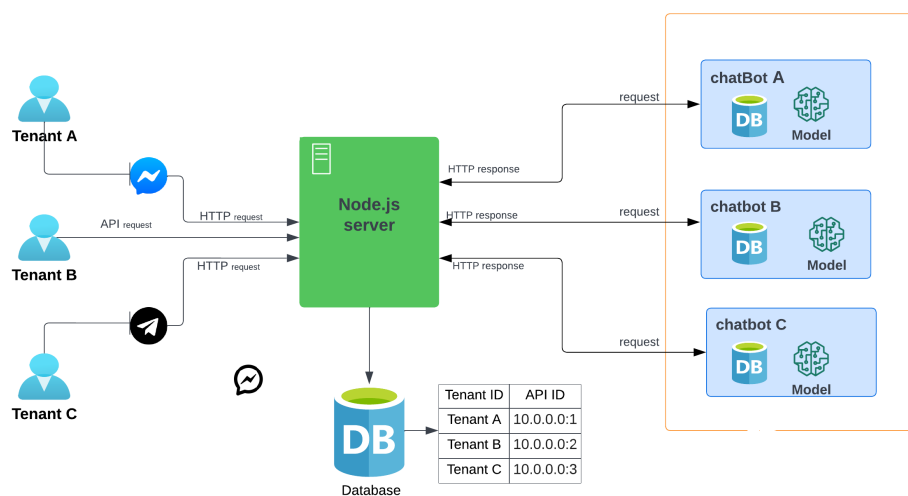
2.2.2 Үйл ажиллагааны загвар



3. СИСТЕМИЙН ЗОХИОМЖ

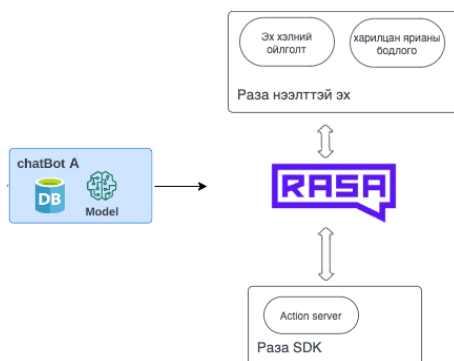
3.1 Системийн ерөнхий архитектур

Доорх зургаар хэрэглэгч/tenant/ болон сервер/server/ хооронд хоёр чиглэлтэй харилцаа холбоог шууд хийдэг тухай харуулсан.



Зураг 3.1: Чатботын ерөнхий архитектур

Чатбот нь rasa фреймворк дээр хөгжүүлэгдэж байна.



Зураг 3.2: Чатботын архитектур

4. ХЭРЭГЖҮҮЛЭЛТ, ҮР ДҮН

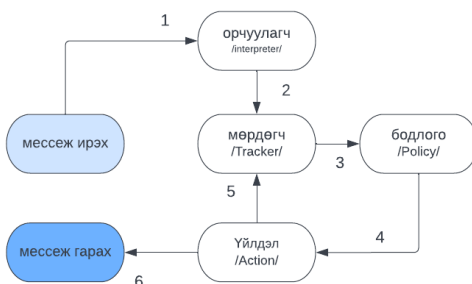
4.1 Rasa фреймворк

Раза бол Python болон эх хэлний ойлголт (NLU) ашиглан хиймэл оюун ухаантай (AI) чатбот бүтээх фреймворк юм. Энэ нь хэрэглэгчдэд загвараа сургах, захиалгат үйлдлүүдийг нэмэх боломжийг олгодог.

4.1.1 Rasa бүтэц

Раса нь хоёр үндсэн бүрэлдэхүүн хэсэгтэй.

- **Раза NLU** - энд хэрэглэгчийн чатны зорилго болон нэгж-объектуудыг буюу бүтэцлэгдсэн мэдээлэл гаргаж авах нь зорилго нь юм. Раза NLU нь хэрэглэгчийн юу хүсэж байгааг ойлгож, зарим утгыг гаргаж авсны дараа мэдээллээ цааш нь илгээдэг. Өгөгдөлдөө ердийн илэрхийлэл(regular expression/regex), хайлтын хүснэгт(look-up table) гэх мэт нэмэлт мэдээллийг нэмж өгснөөр түүний зөв тодорхойлох нарийвчлал нь ихэсдэг байна.
- **Раза Core** - Раза NLU-ийн илгээсэн өгөгдлийг хүлээн авч, хэрэглэгч рүү гаралт болгон илгээх зөв хариултыг олж боловсруулдаг.[7]



Зураг 4.1: Rasa үндсэн үйл ажиллагаа

- 1. Хэрэглэгч чатбот руу мессеж илгээнэ.
- 2. Орчуулагч хэсэгт хүсэлтийг хүлээн авч Раза NLU чухал өгөгдлийг төлөөлөх зарим нэгж-объектуудыг гаргаж аваад ур дүнг мөрдөгчрүү илгээнэ.

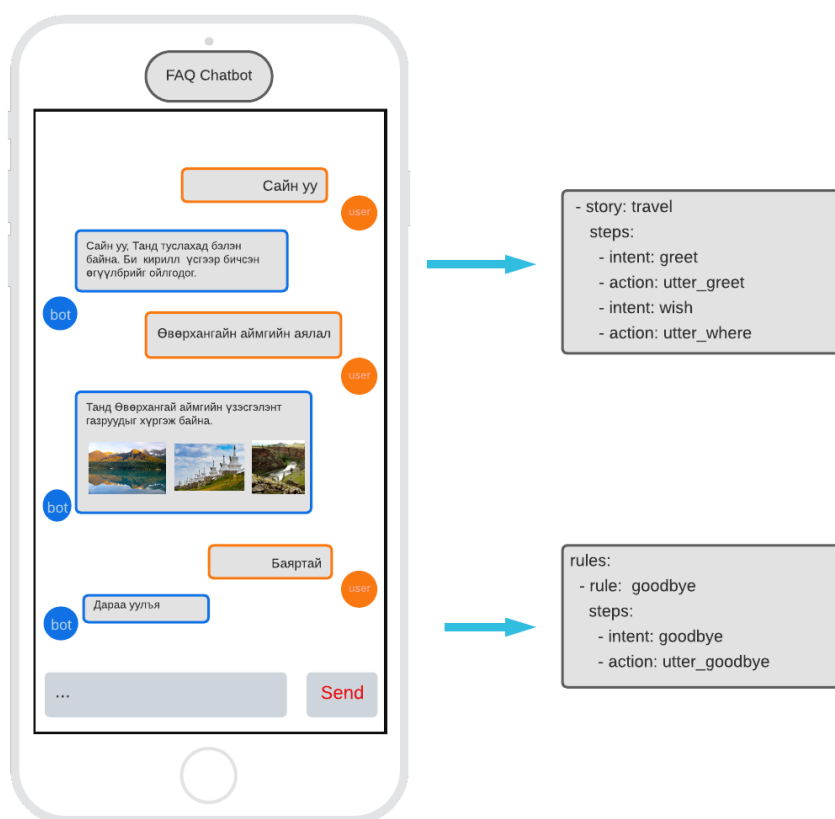
Доорх бусад үйлдлүүд нь Раза Цөм-д агуулагддаг.

- 3. Мөрдөгч (Tracker) нь тухайн моментод өрнөж буй харилцан ярианы түүхийг санах ойд хадгалахад ашиглагддаг. Энд тусгай үйлдэл (custom action) ашиглан өөрийн ботын санах ойд хандах боломжийг олгоно. Мөрдөгч гишүүн өгөгдөл болон гишүүн функцээр дамжуулан өнгөрсөн үйл явдлууд болон өнөөгийн ярианы байр байдлын талаар мэдээлэл авах боломжтой. Жишээ нь, хэрэглэгч нэрээ чатботод хэлж хадгалуулсан бол нэрийг нь хадгалж явцдаа нэрээр нь дуудаж болно. ”Сайн уу, заяа”) Үүний дагуу чатбот нь харилцан яриа хаана байгааг мэдэж байх ёстой бөгөөд энэ нь харилцан ярианы төлөвийг хадгалах болно.
- 4. Бодлого (Policy). Туслах харилцан ярианы алхам бүрт ямар арга хэмжээ авахаа шийдэхийг Бодлого хариуцдаг. Ботдоо сонгож болох өөр өөр бодлого байдаг бөгөөд тохиргоондоо олныг ч оруулах боломжтой. Хамгийн их итгэлтэйгээр таамаглагдсан бодлого нь туслахын дараагийн үйлдлийг шийддэг.
- 5. Энд байгаа үйлдэл (Action), ажлаа хийхийн тулд харилцан ярианы түүхийг нь үзэж юу хийхээ тодорхойлно. Нэг үйлдэл нь хэд хэдэн өөр үйлдлүүдийг өдөөж болно.
- 6. Үйлдэл хийгдэж дуусахад, хэрэглэгч рүү гаралт үүсгэгдэнэ.

4.1.2 Энгийн Раза ботод шаарлагатай файлууд

Раза дээр шинэ project үүсгэхэд дараах файлууд үүсэх бөгөөд эдгээрийг дэлгэрүүлж чатботоо сургана:

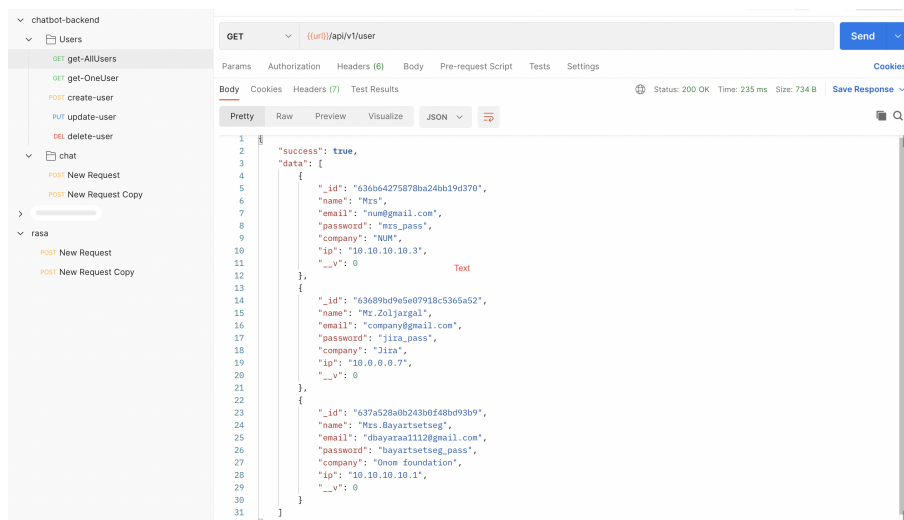
- Domain.yml -Тухайн чатботод хэрэгтэй хариулт илгээхэд шаардлагатай мэдээллүүд байна (form, slot, response ...) .
- NLU.yml - NLU өгөгдөл буюу зорилго, хайлтын хүснэгт (Look-up Table), регекс (regular expression), синоним г.м
- Stories.yml - зохиолуудыг тодорхойлно.
- Rules.yml - дүрмүүдээ тодорхойлно .
- Config.yml - Раза NLU болон Раза Core моделуудийн тохиргоог оруулна.



Зураг 4.2: Зохиол ба дүрмийн жишээ

4.2 Чатботын хэрэгжүүлэлт

Доорхи зурагт чатботонд бүртгэлтэй хэрэглэгчдийн мэдээлэл байна.



Зураг 4.3: Чатбот хэрэглэгчийн endpoint

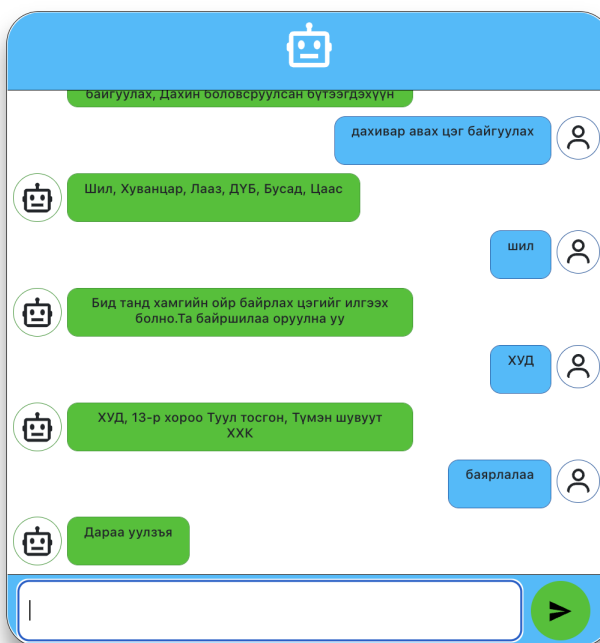
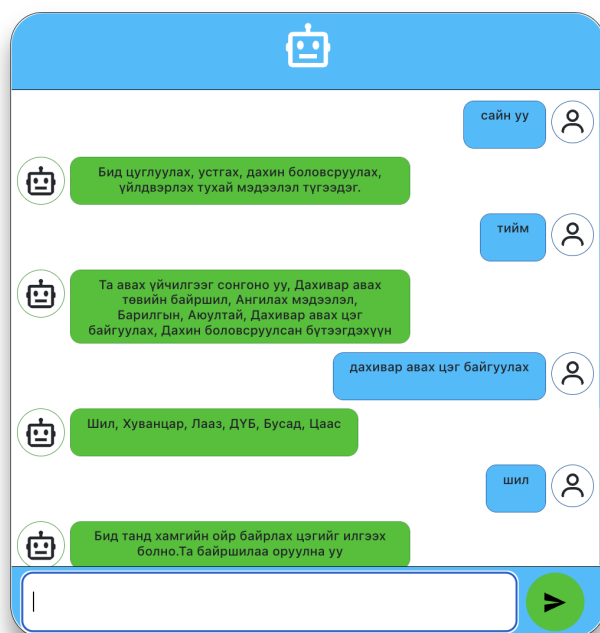
```
Bot loaded. Type a message and press enter (use '/stop' to exit):
Your input -> Сайн байна уу
Сайн уу, Танд туслахад бэлэн байна.
Your input -> Өвөрхангай аялал
Танд Өвөрхангай аймгийн үзэсгэлэнт газруудыг хүргэж байна.
Image: https://i.imgur.com/nGF1K8f.jpg
Your input -> Баяртай
Bye
Your input -> 
```

Зураг 4.4: Туршилтын энгийн ботын үр дүн

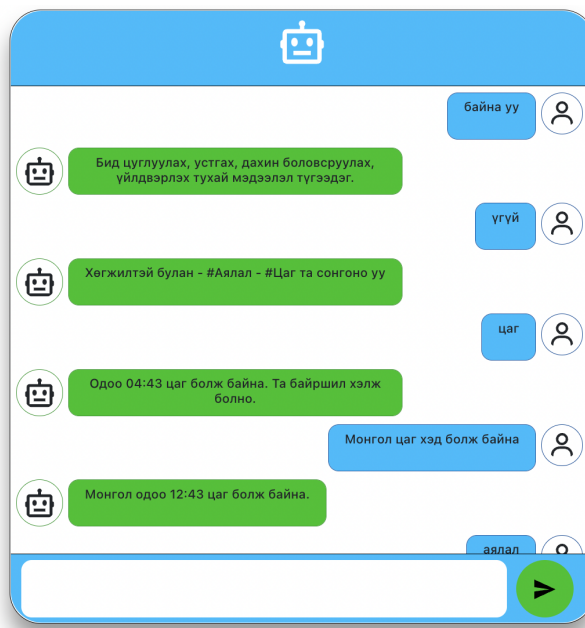
```
Bot loaded. Type a message and press enter (use '/stop' to exit):
Your input -> байна уу
Сайн уу, Танд туслахад бэлэн байна.
Your input -> Цаг London хэд болж байна
вэ?
London одоо 23:30 цаг болж байна.
Your input -> Berlin хэдэн цаг болж байна вэ
Berlin одоо 00:31 цаг болж байна.
Your input -> Amsterdam цаг?
Amsterdam одоо 00:31 цаг болж байна.
Your input -> цаг хэд болж байна вэ
Одоо 23:31 цаг болж байна. Та байрлал хэлж болно.
Your input -> бая
Дараа уулзъя
Your input -> 
```

Зураг 4.5: Туршилтын динамик шинжтэй ботын үр дүн

Демо чатботын хэрэглэгчийн интерфэйс, сервер талыг хэрэгжүүлж кодоо холбож ажиллууллаа.



Зураг 4.6: Туршилтын ботын үр дүн



4.3 Бүлгийн дүгнэлт

Демо бот нь туршилтаар гурван төрлийн мэдээллийг хүргэх чадвартай бөгөөд

- Ботын хэрэглэгч нар дахин боловсруулах хог хаягдал, ангилсан хогоо хаана өгөх түүгээр юу хийж болдог тухай мэдэх боломжтой.

- Аялалын тухай бот

- Хэрэглэгч дэлхийн улс орны цаг хэд болж байгааг мэдэх боломжтой.

Хэрэглэгчээс хүлээн авсан мессеж бүр ID -г агуулах бөгөөд ингэснээр хэрэглэгчид хариу өгөх дахин холбогдох боломжтой.

Дүгнэлт

Мэдээлэл хайлтын явцыг хялбарчлах шийдэл бол чатбот юм. Энэхүү судалгааны ажлын хүрээнд мэдээлэл солилцох боломжтой монгол дээрхи чатботын бичил хэсгүүдийг хөгжүүллээ.

Үүнд:

- Хаягдал хог ангилах, дахивар хаягдлын талаар мэдээлэл түгээх мөн хэрэглэгч байршилаа оруулан өөрт хамгийн ойр байрлах үйлчилгээний төвийг авах түүнд хариу үйлдэл үзүүлэх бот

- Дэлхийн улсуудын хэдэн цаг болж байгаад хариулах бот

- Улс орны үзэсгэлэнт газруудыг харуулах асуултанд хариулах боломжтой энгийн тогтмол асуулт хариулттай болон динамик олон хариулттай демо бот тус бүрийг боловсруулсан.

Цаашид чатбот нь асуултанд зөв хариулах баялаг боловсруулалт, зорилго, нэгж-объектыг зөв таних, харилцан яриаг илүү ойлгомжтой, ухаалаг болгох хэсгүүдэд сайжруулалт хийх хэрэгцээтэй байна.

Bibliography

- [1] <https://medium.com/@sudheer.sandu/multi-tenant-application-68c11cc68929>
- [2] [https://manychat.com/blog/chatbot/?utm_source=google&utm_medium=pmax&utm_campaign=\[PMAx\] \[2022\] \[InMarket/Affinity\] \[C:US/CA\]&utm_content=consideration&gclid=EAIaIQobChMI4de-zc7o-gIVzDmrCh14pAYkEAAAYAiAAEgK8-_D_BwE](https://manychat.com/blog/chatbot/?utm_source=google&utm_medium=pmax&utm_campaign=[PMAx] [2022] [InMarket/Affinity] [C:US/CA]&utm_content=consideration&gclid=EAIaIQobChMI4de-zc7o-gIVzDmrCh14pAYkEAAAYAiAAEgK8-_D_BwE)
- [3] <https://www.amazon.com/Node-js-Design-Patterns-server-side-applications-ebook/dp/B01D8HIIFU> Mario Casciaro, Luciano Mammino, Node.js Design Patterns - Second Edition. 2016
- [4] <https://aws.amazon.com/docker/>
- [5] MongoDBvsMySQL: A Comparative Study on Databases, <https://www.simform.com/mongodb-vs-mysql-databases>
- [6] https://en.wikipedia.org/wiki/Natural_language_processing
- [7] <https://www.geeksforgeeks.org/chatbots-using-python-and-rasa/?tab=article>

A. КОДЫН ХЭРЭГЖҮҮЛЭЛТ

A.1 Action хэсгийн код

```
1 from typing import Any, Text, Dict, List
2 import arrow
3 import dateparser
4 from rasa_sdk import Action, Tracker
5 from rasa_sdk.executor import CollectingDispatcher
6 from rasa_sdk.events import SlotSet
7
8 city_db = {
9     'london': 'Europe/Dublin',
10    'amsterdam': 'Europe/Amsterdam',
11    'Berlin': 'Europe/Berlin',
12    'Mongolia': 'Asia/Ulaanbaatar',
13    'San Francisco': 'America/Los_Angeles',
14    'Paris': 'Europe/Paris',
15    'New York': 'America/New_York'
16 }
17
18 class ActionCustomPlace(Action):
19     def name(self) -> Text:
20         return "action_time"
21
22     def run(self, dispatcher: CollectingDispatcher,
23            tracker: Tracker,
```

```

24         domain: Dict[Text, Any]) -> List[Dict[Text, Any]]:
25         current_place = next(tracker.get_latest_entity_values("place"),
26                               None)
27         utc = arrow.utcnow()
28         if not current_place:
29             msg = f" Odoo {utc.format('HH:mm')} tsag болж байна Та
30                 bairshil helj болно."
31             dispatcher.utter_message(text=msg)
32             return []
33
34         tz_string = city_db.get(current_place, None)
35         if not tz_string:
36             msg = f"bi uuniig {current_place} tanisangui Zuw bichne uu
37                 "
38             dispatcher.utter_message(text=msg)
39             return []
40
41         msg = f"{current_place} odoo {utc.to(city_db[current_place]).
42               format('HH:mm')} tsag болж байна."
43         dispatcher.utter_message(text=msg)
44         return []

```

A.2 Харилцан ярианы түүх хэсэг

```

1 stories:
2 #bot1 aylaliin tuhai medelel haruuldag.
3     - story: travel
4         steps:
5             - intent: greet

```

```

6         - action: utter_greet1
7         - intent: wish
8         - action: utter_where
9 #bot2 delhiin ulsiin tsag haruuldag.
10    - story: information
11      steps:
12        - intent: greet
13        - action: utter_greet1
14        - intent: what_time
15        - action: action_time
16 #bot3 haywal hog angilwal mungu
17    - story: my
18      steps:
19        - intent: greet
20        - action: utter_greet1
21        - intent: response
22        - action: utter_choose
23        - intent: my_choose
24        - action: utter_location
25        - intent: my_location
26        - action: utter_loc
27 #bot4 Dahiwar hgzygdal baiguulah
28    - story: my_build
29      steps:
30        - intent: greet
31        - action: utter_greet1
32        - intent: response
33        - action: utter_menu

```

```

34     - intent: my_menu
35     - action: utter_choose
36     - intent: my_choose
37     - action: utter_location
38     - intent: my_location
39     - action: utter_loc
40     - intent: my_next
41     - action: utter_method
42     - intent: my_method
43 - story: my_info
44   steps:
45     - intent: greet
46     - action: utter_greet1
47     - intent: my_next
48     - action: utter_nomenu
49     - intent: response
50     - action: utter_where
51     - action: action_time

```

A.3 Хайх lookup жагсаалт

```

1 #bot2
2   - intent: what_time
3   examples: |
4     - tsag
5     - tsag hed bolj baina we
6     - [Mongolia](place)
7     - [Berlin](place)

```

```

8         - [London](place)          ?
9         - [Amsterdam](place)      ?
10        - (place)
11    - lookup: place
12    examples: |
13        - london
14        - amsterdam
15        - Berlin
16        - Mongolia
17        - San Francisco
18        - New york

```

A.4 Ботын хариулах хэсгийн код

```

1 version: "3.0"
2 intents:
3     - greet
4     - wish
5     - what_time
6     - place
7     - affirm
8     - my_menu
9     - my_waste
10    - my_location
11    - deny
12    - goodbye
13    - my_choose
14 entities:

```

```

15   - place
16 actions:
17   - action_time
18 responses:
19   utter_where:
20     - text: "Tand Uvurkhangai aimgiin uzesgelent gazruudiig hurgej
           baina                      ."
21   utter_greet1:
22     - text: "Sain baina uu Ta dot-d holbogdloo Tawtai moril. Ta
           guilchilgee haij baina uu"
23   utter_nomenu:
24     - text: " Hugjiltei bulan #Aylal #Tsag Ta songono uu"
25   utter_choose:
26     - text: " # Shil, # Huwantsar, # Laaz # Busad, # Tsaas"
27   utter_menu:
28     - text: " Ta avakh uilchilgeeg songono uu. #Dahiwar awah tuwiin
           bairshil, #Dahiwar khug, #Dahiwar awah tseg baiguulah, #Dahin
           bolowsruulsan buteegdehuun"
29
30   utter_location:
31     - text: "Bid tand hamgiin oir bairlah tsegiig ilgeekh bolno. Ta
           bairshilaa oruulna uu"
32   utter_phone_method:
33     - text: "Tantai holboo barikh dugaar hed we"
34   utter_user_question:
35     - text: "Uur uilchilgee awakh uu"
36   utter_submit:
37     - text: "Bugdiig huleen awlaa."

```

```

38 utter_slots_values:
39   - text: "Bid holbogdokh bolno. Tanii medeelel:\n
40     - user_location: {user_location}\n
41     - user_phone: {user_phone}\n
42     - user_question: {user_question}"
43 utter_loc:
44   - text: "HUD 13-r khoroo Tuul tosgon Tumen shuwuut XXK"
45 utter_goodbye:
46   - text: "Daraa uulziy"
47
48 session_config:
49   session_expiration_time: 60
50   carry_over_slots_to_new_session: true

```

A.5 Ботын дүрэм боловсруулалт

```

1 recipe: default.v1
2 language: mn
3 pipeline:
4   # hooson zaigaar tusgaarlah bolomjtoi nlu
5   #hooson zaigaar tusgaarlagdsan temdegt buriin daraalald temdegt
6   uusgene
7   - name: WhitespaceTokenizer
8   # negj zadlah, intent angilah
9   - name: RegexFeaturizer
10  - name: LexicalSyntacticFeaturizer
11  # ugeer ilerhiilne
12  - name: CountVectorsFeaturizer

```

```

12 - name: CountVectorsFeaturizer
13     analyzer: char_wb
14     min_ngram: 1
15     max_ngram: 3
16 # negj zadlah, intent angilah tanihad zoriulagdsan arhitektur
17 - name: DIETClassifier
18     epochs: 100
19     constrain_similarities: true
20 - name: EntitySynonymMapper
21 - name: ResponseSelector
22     epochs: 100
23     constrain_similarities: true
24 - name: FallbackClassifier
25     threshold: 0.3
26     ambiguity_threshold: 0.1
27
28 - name: RegexEntityExtractor
29     use_lookup_tables: true
30 policies:
31 # stories ni odoogiin yriatai taarch bnuug shalgan
32 - name: MemoizationPolicy
33 # yriani hesgiig zohitsuuldag bodlogo
34 - name: RulePolicy
35 - name: UnexpectedIntentPolicy
36     max_history: 5
37     epochs: 100
38 - name: TEDPolicy
39     max_history: 5

```



```
40     epochs: 100
41     constrain_similarities: true
42     # durmiin bodlogo togtson durmeer yriag zohitsuulna
43 - name: RulePolicy
```