

# LSBU Course Finder Chatbot

Module Title: Future Internet Technologies (CSI\_7\_FIT)

## Team Members:

Syed Jafrul Husen (4037556)  
Shourob Ahmed Khan (4108471)  
Syed Ishmam Mahmud (4103190)  
Md Mehedhi Hasan (3626236)

# What is chatbot?



Chatbot is a software application that allow humans to interact with digital device(text or speech)

# LSBU course finder chatbot



A user, ideally student, who want to know details about a specific course it will gives proper response in chatbot as a text format.

# Use case scenerio

- New student
- Parents/ guardian
- University admission team



# Tool and Technology

- Rasa 2.0
- Python 3.8.10
- Visual studio as a code editor
- Slack for deployment

# Development model

To develop this project, we followed **waterfall model**.

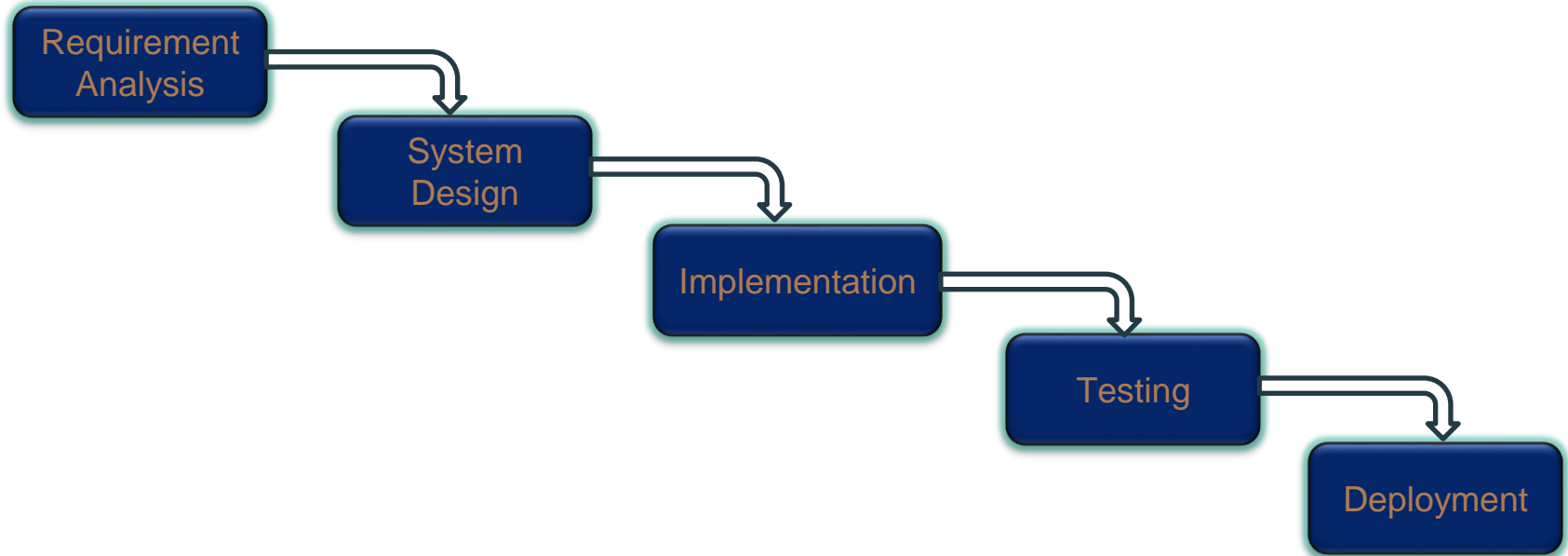


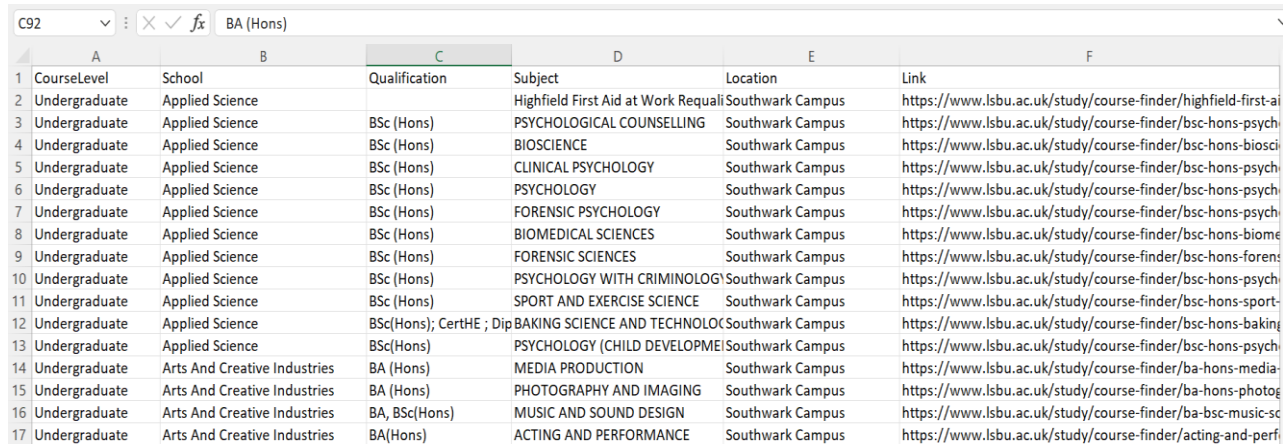
Fig 1: Waterfall Model

# Why used waterfall model?

- Easy to use and manage due to rigid model
- Works well where requirement are specified and the project is small
- Easy to arrange task
- Clearly defined stages

# Data collection

- LSBU website
- Collected details information about 150 subjects
- Data collected into csv file (comma delimited)



	A	B	C	D	E	F
1	CourseLevel	School	Qualification	Subject	Location	Link
2	Undergraduate	Applied Science		Highfield First Aid at Work Requali	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/highfield-first-ai">https://www.lsbu.ac.uk/study/course-finder/highfield-first-ai</a>
3	Undergraduate	Applied Science	BSc (Hons)	PSYCHOLOGICAL COUNSELLING	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych</a>
4	Undergraduate	Applied Science	BSc (Hons)	BIOSCIENCE	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-biosci">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-biosci</a>
5	Undergraduate	Applied Science	BSc (Hons)	CLINICAL PSYCHOLOGY	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych</a>
6	Undergraduate	Applied Science	BSc (Hons)	PSYCHOLOGY	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych</a>
7	Undergraduate	Applied Science	BSc (Hons)	FORENSIC PSYCHOLOGY	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych</a>
8	Undergraduate	Applied Science	BSc (Hons)	BIOMEDICAL SCIENCES	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-biome">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-biome</a>
9	Undergraduate	Applied Science	BSc (Hons)	FORENSIC SCIENCES	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-forens">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-forens</a>
10	Undergraduate	Applied Science	BSc (Hons)	PSYCHOLOGY WITH CRIMINOLOG	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych</a>
11	Undergraduate	Applied Science	BSc (Hons)	SPORT AND EXERCISE SCIENCE	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-sport">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-sport</a>
12	Undergraduate	Applied Science	BSc(Hons); CertHE ; Dip	BAKING SCIENCE AND TECHNOLO	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-baking">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-baking</a>
13	Undergraduate	Applied Science	BSc(Hons)	PSYCHOLOGY (CHILD DEVELOPME	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych">https://www.lsbu.ac.uk/study/course-finder/bsc-hons-psych</a>
14	Undergraduate	Arts And Creative Industries	BA (Hons)	MEDIA PRODUCTION	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/ba-hons-media">https://www.lsbu.ac.uk/study/course-finder/ba-hons-media</a>
15	Undergraduate	Arts And Creative Industries	BA (Hons)	PHOTOGRAPHY AND IMAGING	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/ba-hons-photog">https://www.lsbu.ac.uk/study/course-finder/ba-hons-photog</a>
16	Undergraduate	Arts And Creative Industries	BA, BSc(Hons)	MUSIC AND SOUND DESIGN	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/ba-bsc-music-sc">https://www.lsbu.ac.uk/study/course-finder/ba-bsc-music-sc</a>
17	Undergraduate	Arts And Creative Industries	BA(Hons)	ACTING AND PERFORMANCE	Southwark Campus	<a href="https://www.lsbu.ac.uk/study/course-finder/acting-and-perf">https://www.lsbu.ac.uk/study/course-finder/acting-and-perf</a>

Fig 2: Sample dataset



# Technical details

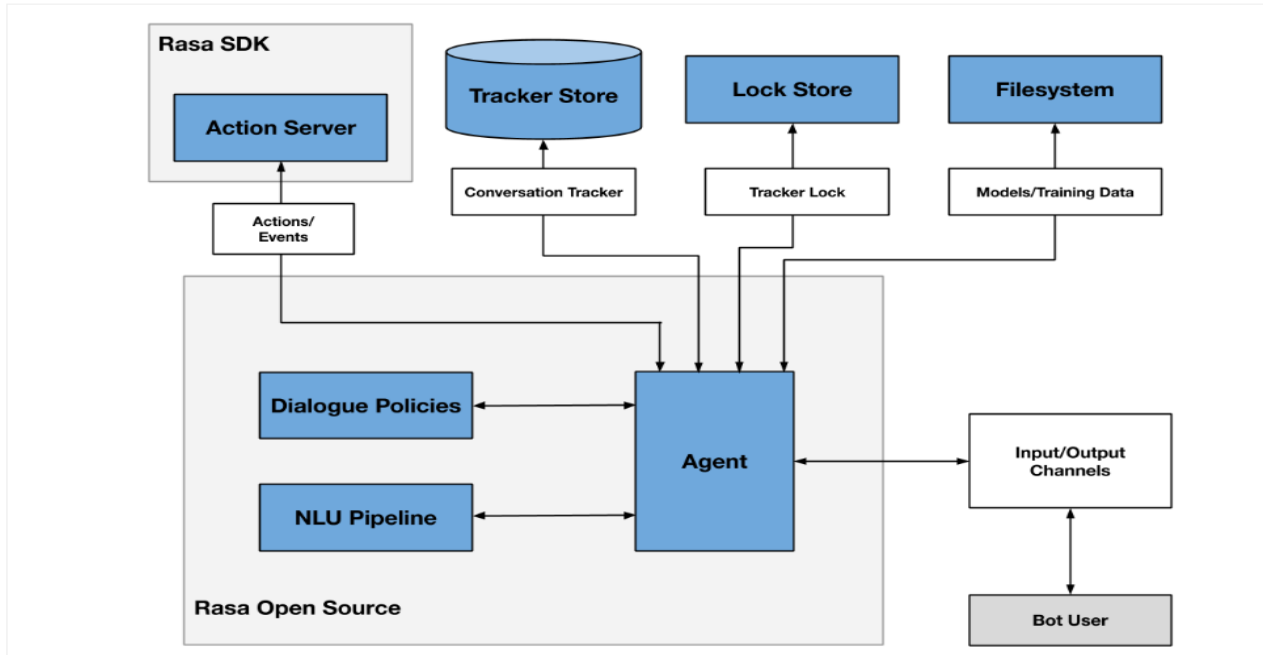


Fig 3: RASA architecture [2]

# Pipeline

```
pipeline:
- name: WhitespaceTokenizer
- name: RegexFeaturizer
- name: LexicalSyntacticFeaturizer
- name: CountVectorsFeaturizer
- name: CountVectorsFeaturizer
  analyzer: char_wb
  min_ngram: 1
  max_ngram: 4
- name: DIETClassifier
  epochs: 100
- name: EntitySynonymMapper
- name: ResponseSelector
  epochs: 100
- name: FallbackClassifier
  threshold: 0.3
  ambiguity_threshold: 0.1
```

Fig 4: Pipeline

# Policy

```
policies:  
  - name: MemoizationPolicy  
  - name: TEDPolicy  
    max_history: 5  
    epochs: 100  
  - name: RulePolicy
```

Fig 5: Policy

## Technical details cont...

- Entity:
  - Course\_level
  - Course
- Custom Action:
  - Action\_course\_search

# Spelling corrector

- Algorithm: Peter norvigs
- Used Baye's theorem:  $\operatorname{argmax}_{c \in candidates} P(c) P(w|c) / P(w)$
- Four parts of this function:
  - Selection mechanism:  $\operatorname{argmax}$
  - Candidate model:  $c \in candidates$
  - Language model:  $P(c)$
  - Error Model:  $P(w/c)$

# Project Demo

# Performance analysis

- Accuracy of spell corrector algorithm: 80% to 90%

```
2022-05-10 16:01:05 INFO rasa.nlu.test - Intent evaluation results:
2022-05-10 16:01:05 INFO rasa.nlu.test - Intent Evaluation: Only considering those 170 examples that have a defined
intent out of 170 examples.
2022-05-10 16:01:05 INFO rasa.nlu.test - Classification report saved to results\intent_report.json.
2022-05-10 16:01:05 INFO rasa.nlu.test - Incorrect intent predictions saved to results\intent_errors.json.
2022-05-10 16:01:05 INFO rasa.utils.plotting - Confusion matrix, without normalization:
[[19  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  5  0  0  0  0  0  0  0  1  0  2  0]
 [ 0  0  6  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  8  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  5  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  7  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 15  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 22  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 30  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 16  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 16  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 18]]
2022-05-10 16:01:07 INFO rasa.nlu.test - Entity evaluation results:
2022-05-10 16:01:07 INFO rasa.nlu.test - Evaluation for entity extractor: DIETClassifier
2022-05-10 16:01:07 INFO rasa.nlu.test - Classification report for 'DIETClassifier' saved to 'results\DIETClassifier_report.json'.
2022-05-10 16:01:07 INFO rasa.nlu.test - Your model predicted all entities successfully.
2022-05-10 16:01:07 INFO rasa.utils.plotting - Confusion matrix, without normalization:
[[ 14  0  0]
 [  0  5  0]
 [  0  0 423]]
```

Fig 6: Confusion Matrix

# Testing

- Unit test:
  - Test separately python code as in rasa it take time to run and hard to integrate large code
    - Every unit of code tested separately (whitespace remove, edit distance calculation, custom actions, entity, slots, domain etc.)
- Integration test:
  - Evaluate how a piece of code work with the project
- Functional testing:
  - Check How the entire system works together



# Teamwork

Task	Details
Requirement Analysis	As a team we collected the requirement and do analysis to make the project effective.
Data collection	Ishmam and Shourob collected data.
Create story	Jafrul and Mehedi created rasa story
System Design	Mehedi and Shourob did the requirement specification, specify the hardware and software requirements and architecture.
Implementation	We together implement the project and did coding using rasa framework. Jafrul implemented spelling correction and rasa custom actions.
Testing and troubleshooting	Ishmam and Jafrul did testing and troubleshooting.
Deployment	We together and made final discussion and deployed it using slack.

# Troubleshooting

```
Attempting uninstall: ruamel.yaml
Found existing installation: ruamel-yaml 0.15.87
ERROR: Cannot uninstall 'ruamel-yaml'. It is a distutils installed project and thus we cannot accurately determine which files belong to it which would lead to only a partial uninstall.
```

```
cloudpickle, fsspec, looliz, pyyaml, workx, typeguard, tensorflow-addons,
rocketchat-API, rasa
ERROR: After October 2020 you may experience errors when installing or updating packages. This is because pip will change
the way that it resolves dependency conflicts.
```

```
We recommend you use --use-feature=2020-resolver to test your packages with the new resolver before it becomes the default.
```

```
pyqt4 4.1.5 requires PyQt5<5.13; python_version >= "3", which is not installed.
pyqt4 4.1.5 requires PyQtwebengine<6.13; python_version >= "3", which is not installed.
sanic 21.12.1 requires multidict<6.0,>=5.0, but you'll have multidict 6.0.2 which is incompatible.
sanic-jwt 1.7.0 requires pyjwt==2.1.0, but you'll have pyjwt 2.3.0 which is incompatible.
rasa 3.1.0 requires ruamel.yaml<0.17.0,>=0.16.5, but you'll have ruamel-yaml 0.17.21 which is incompatible.
```

```
Collecting PyQt5
  Downloading PyQt5-5.15.6-cp36-abi3-win_amd64.whl (6.7 MB)
    | 6.7 MB 252 kB/s
Collecting PyQt5-sip<13,>=12.8
  Downloading PyQt5_sip-12.9.1-cp38-cp38-win_amd64.whl (77 kB)
    | 77 kB 771 kB/s
Collecting PyQt5-Qt5>=5.15.2
  Downloading PyQt5_Qt5-5.15.2-py3-none-win_amd64.whl (50.1 MB)
    | 36.4 MB 233 kB/s eta 0:00:59
```

```
ERROR: THESE PACKAGES DO NOT MATCH THE HASHES FROM THE REQUIREMENTS FILE. If you have updated the package versions, please
update the hashes. Otherwise, examine the package contents carefully; someone may have tampered with them.
```

```
PyQt5-Qt5>=5.15.2 from https://files.pythonhosted.org/packages/37/07/5d3b222b924fa2ed4c2488925155cd0b03fd5d09ee1cfcf7c553c11c9f66/PyQt5_Qt5-5.15.2-py3-none-win_amd64.whl#sha256=750b78e4dba6bdf1607febedc08738e318ea09e9b10aea9ff0d73073f11f6962
(from PyQt5):
Expected sha256 750b78e4dba6bdf1607febedc08738e318ea09e9b10aea9ff0d73073f11f6962
Got 02614a6082d93281c6d080802854f6f0d613f8eae0d3bbdbf6a37f83cb103d
```

```
cheduler, tarsafe, rasa
ERROR: Could not install packages due to an EnvironmentError: [WinError 5] Access is denied: 'C:\\Users\\Unique-Pc\\anaconda3\\Lib\\site-packages\\_cffi_backend.cp38-win_amd64.pyd'
Consider using the '--user' option or check the permissions.
```

**Solved:**  
pip install rasa --ignore-installed ruamel.yaml --user

Solved the raised another error and changed the device

Fig 7: Screenshot of some error during implementation of the project

# Troubleshooting cont...

- Started coding into rasa version 3.0 then migrate the version and change the coding format because get to know rasa 3.0 don't support deployment using rasa which is our future work.

```
from the `scipy.sparse` namespace, the `scipy.sparse.base` namespace is deprecated.  
from scipy.sparse.base import spmatrix  
YamlValidationException: Failed to validate 'C:\Users\Unique-Pc\OneDrive\Desktop\chatbotInternet\domain.yml'. Please make sure the file is  
correct and all mandatory parameters are specified. Here are the errors found during validation:  
in C:\Users\Unique-Pc\OneDrive\Desktop\chatbotInternet\domain.yml:22:  
  Cannot find required key 'mappings'. Path: '/slots/location'
```

Fig 8: Domain.yml file error

# Challenges, Limitation and Future Work

- Challenging task to make the project workable as there has many dependencies between python and rasa.
- Need very clear and sophisticated data and story to satisfy user
- Future work, Improve the story and increase the dataset
- Future work, make the application more dynamic and interactive with human.

# References

- [1] [https://www.lsbu.ac.uk/study/course-finder?collection=LSBU\\_Courses\\_Meta&query=!nullsearch&start\\_rank=1&sort=relevance&f.Level\\_new|courseLevel=undergraduate](https://www.lsbu.ac.uk/study/course-finder?collection=LSBU_Courses_Meta&query=!nullsearch&start_rank=1&sort=relevance&f.Level_new|courseLevel=undergraduate)
- [2] <https://rasa.com/docs/rasa/2.x/arch-overview>
- [3] <https://norvig.com/spell-correct.html>
- [4] <https://rasa.com/blog/unpacking-the-ted-policy-in-rasa-open-source/>

**Thank You**