# Design

Based upon our previous section detailing the specific problems we would like to address, this section will contain the design and implementation phases of the project based upon our previously identified objectives. The design phase will focus on detailing the individual modules and how they can interconnect to form a working solution. The implementation phase will look at the final product, as well as any issues encountered along the way that may have differed from the plan.

The general design ethic for the project is for it to be modular. This allows for the replacement of entire services with others as long as the data output remains consistent with other modules. For example, replacing Microsoft’s LUIS natural language processing service with Facebook’s Wit.ai would not require an entire system rewrite, but would only require the replacement of the Natural Language Processing module.

As an addition to the modular design ethic, the project is intended to be a framework that can be refined, worked on and added to over time. It is not intended to be a complete monitoring system for a network – rather it is a framework with a various proof of concept and example uses pre-programmed.

It is intended that the application be semi-autonomous in nature, able to be extremely low maintenance once configured. The application will have two operational modes:

**Interactive mode** is a state where the user leads the conversation. A user will query the app, which will cause the app to produce a result and serve it back to the user.

**Monitor mode** is the applications default state, where it is analysing incoming information from sources such as servers and comparing it to known thresholds. If a threshold is met, a user responsible for that threshold is informed – The application is then in interactive mode once a user replies.

### Data Sources

For our testing we will only be taking data from a production Linux Ubuntu operating system. This will include the following data sources:

**Apache HTTP Server Access Logs** show records of pages served and files loaded by the webserver. This can be valuable information if formatted correctly. This information will be used to generate statistics.

**Apache HTTP Server Error Logs** show records of all error conditions reported by the server, which in some cases will need urgent attention.

**Authorization Logs** track usage of authorization systems such as *sudo* and remote logins over SSH. These can be used for both statistical purposes and for showing login attempts.

**Login Failures Log** is designed to be non-human-readable, and contains all login failures. This may be better for machine parsing than the authorization logs.

**Last Logins Log** is also non-human-readable, and shows the last login of users.[1]

However, there is no reason why extra modules could be added with more commands and to accept more sources.

### Other Commands

Remote administration means being away from tools, so adding tools that are often in an administrators’ arsenal to the application can be very useful in troubleshooting. Having remote access to a machine inside the network means that certain commands can be piped to the user:

**Ping** can be used to ping both internal and external resources, telling the user if that resource is online and what its connection is like.

**Traceroute** shows the user the route packets are taking to a server.

**Nslookup** allows a user to look the DNS information of a resource either internal or external as reported by their DNS server.

**SSH** allows an interactive shell to another computer. This may or may not be possible with our planned setup and will have to be tested during the implementation.

## Security and Other Considerations

Security is a major point in this project, as this application may have access to logs or passwords for servers. To address this issue, we will be using the security offered by the various chat services by associating access to the bot with access to instant messaging services.

In our case, BotBuilder provides a user ID along with every message. We can compare this user ID to a hardcoded whitelist that allows a user access to functions in the bot if they are whitelisted. A user should take steps such as two-factor authentication on their IM account in order to protect it, and by extension, any data the project may be able to give out.

It would be possible to add more security considerations to the modules such as passwords, but this did not seem necessary for this example framework.

## Modules

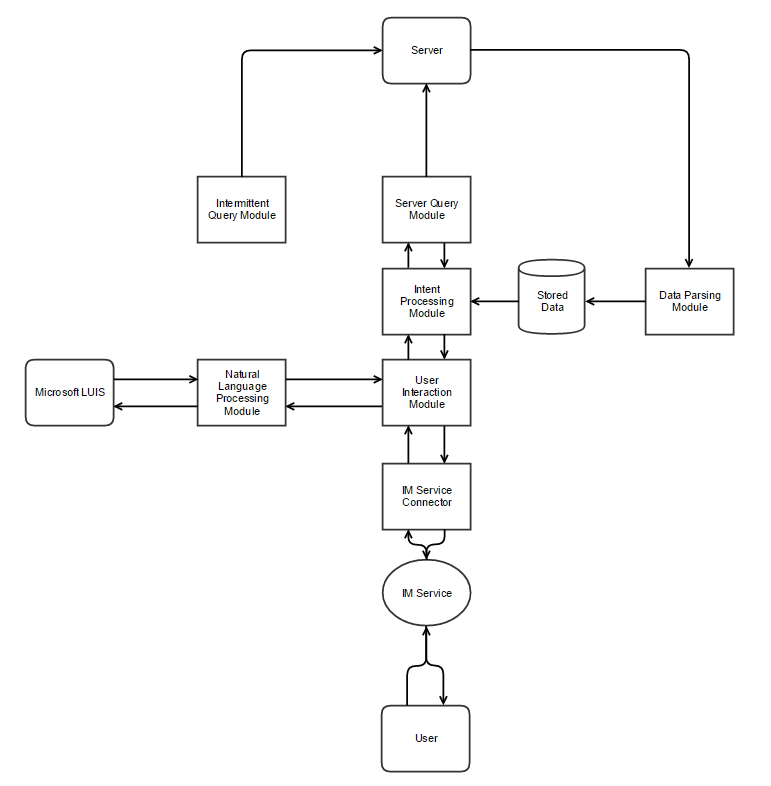


Figure 1 - Technical Prototype Flow Chart

Our initial plan is shown in Figure 1 - Technical Prototype Flow Chart, and shows each individual module. The following section will go into detail of each module and discuss its uses, as well as how they will interact or be swapped out as necessary.

### IM Service Connector

This module connects the User Interaction module to the IM service. This simply manages the connection to the instant messaging server, be it IRC, Jabber, Telegram or any other. As we already know the Microsoft Bot Builder[2] acts as a proxy between many IM services and our application, some of the work here is done for us. However, in the case that a client wishes to use a different service, our User Interaction Module must be fed data in the same format.

Bot Builder does the work of connecting to the IM service and then forwarding user data from the IM to our API endpoint. Following Bot Builder conventions, we will use their library to open a Restify[3] web service on our server, which will receive the data and handle replies to the Bot Builder server. We can expect the data to be very similar irrelevant of what IM service the user is using. This data is handed directly off to the NLP service without modification, and contains the session information such as the user’s ID, what conversation it is part of, other data required for a reply, and the users message.

Bot Builder was chosen due to it’s also modular nature, it allows for different NLP back-ends to be replaced with relative ease as well as not using one entirely. It also allows for many different IM services to use the same bot back-end, so as not to limit the user in their choices and it can be easily integrated into existing company instant messaging systems.

### Natural Language Processing Module

This module takes the formatted IM messages and sends them to an NLP service, and formats the information to be understood by the User Interaction Module. For the sake of our testing, this module will be tuned to Microsoft LUIS[4] running on an Azure instance.

Intents are what the user ‘intends’ to do. These intents, as well as what triggers them are defined in the LUIS NLP website, and will be planned later.

At program startup, the NLP module queries the User Interaction(UI) Module to get a list of known intents. The NLP module then generates templates to allow it to format the information in a way the UIM understands. No matter what service the NLP module connects to, it will format the information the same way. This generation of functions means that all data coming to the UI module will be in a constant format irrelevant of the intent, and as long as the NLP module sends that data in the same way it is irrelevant which NLP service is in use. We preferred this as it meant significantly less hard-coding of intent functions within the NLP module, making it easier to keep consistency in the event that the NLP module is replaced with one for a different service.

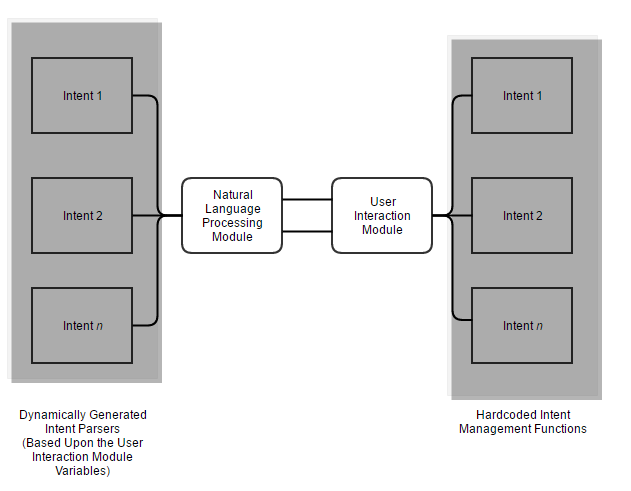


Figure 2 - Showing the relationship between the NLP Module and UI Module

Shown in Figure 2, the intent management functions in the UI Module are hard-coded. The NLP Module then dynamically generates the functions required to handle these incoming intents. The functions here also handle the reply to the user, so it is important to make good use of callbacks between the two modules so the session data is not lost.

Upon receiving a message, the NLP Module will receive the data and immediately send it to our NLP Service (LUIS) to be analysed, which will then be returned and associated with a function related to the intent that the NLP Service deemed the user was trying to activate. The function will then strip the arguments from the formatted message provided by the NLP Service into a more generic format, shown in Figure 3.

This generic format will contain the values of all arguments in the users message. For example, if the user were to type “Show me the log files for Apache on the Ubuntu server”. The intent would be ‘log’, and the two arguments would be ‘Apache’ and ‘Ubuntu’. The NLP Module sends this data, along with the session data in case it is needed later, plus a callback to the UI Module to be handled. The callback, when triggered by the UI module, will accept three arguments. The first argument will be either a string or array – If a string is sent, it is sent directly to the user. If an array is sent, it will be looped and each part of the array will be sent as a separate message to the user. The second argument specifies to the function that the dialog should be handed off to another function that handles multi-message dialogs.

{

**"application1"**:"value1",

**"argument1"**:"value2",

**"server1"**:"value3",

**"argumentN"**:"value4"

}

Figure 3 - Generic argument format

The function that handles multi-message dialogs will be hardcoded, but still will call the intent handler in the UI module for instructions with a specified keyword ‘Convo’, which states that the conversation has entered a multi-message dialog state. This function will loop, receiving data based on how many times it has looped from the UI module with what to send to the user. The UI module may call back with a variable dictating the type of message to send the user, either a string, a prompt or an exit. A string will simply send a message to the user and loop again. An exit will exit the dialog.

However, a prompt is a special situation. It first prompts the user with a question, and then hands off to another handler function using Bot Builder’s Waterfall method, which allows for conversations to be handled differently depending on the flow of a conversation. This final function then adds the users response to an array, and loops back to the original dialog function in order to call and send that data to the UI Module for continued processing.

### User Interaction(UI) Module

This module transports incoming information from the Natural Language Processing module and forwards formatted data of the users’ intent to the Intent Processing Module. This module also formats data intended for the user. In some cases, Intent Processing isn’t necessary and a simple reply can be managed by the User Interaction Module. This module is entirely based on the ‘Interactive Mode’ idea, where it is only in use when the user is directly using the application.

This module can be left static between changes of the NLP and IM modules, as it relies on those modules handling formatting and simply has a list of known intents and how it reacts to them.

The UI Module acts as the middleware between the raw data of the NLP Module and the Intent Processing module. It contains a hard-coded list of intents which should be identical to those specified and trained on the NLP service website, which will act as the basis for the NLP modules functions. As by this point the intent has already been figured out, and the arguments have been correctly parsed, the UI module is where the real human-written responses start to take place.

In the UI module will be an intent handler, containing a function for each known intent plus conversation function for any intents requiring question and answer type conversations. Each function will accept 2-3 arguments, the first being the callback to the NLP Module that sends messages to the user, the second being any arguments that came from the user and formatted by the NLP Module, and finally the session variable which contains all the user data. These functions will then either then call the callback directly to reply to the user with a predefined variable with minor processing such as picking out the user ID and sending it back, or will pass on data to the Intent Processing module with another callback to return back when finished. The flow of data is shown in Figure 4.

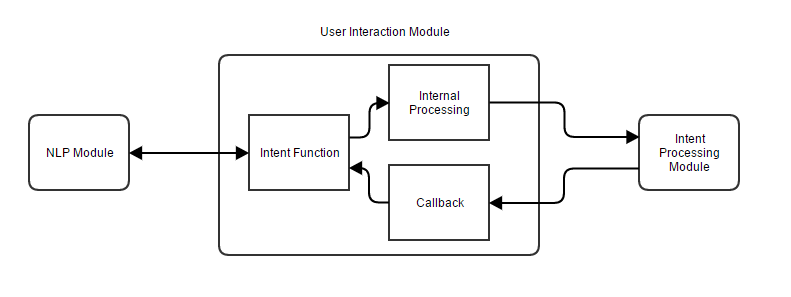


Figure 4 - User interaction module data flow

As an example of a simple intent, such as the ‘help’ intent, the UI module would ignore any arguments and simply call the callback with a simple, hardcoded reply. In the case of an intent requiring responses such as the ‘SSH’ intent, the function would immediately call back to the NLP module requesting a change to the conversation handling function. The conversation handler then loops, incrementing a counter that calls the SSHConvo intent, getting a different prompt or text response each time until all the required information is obtained. This seems like a reliable method for dynamically generating conversations without a complex hardcoding system.

To prove this framework is viable, we will implement several simple responses and 1-2 complex responses. The list of planned user interactions, and therefore intents, are:

* None: Will respond that no intent has been detected.
* Version: Will respond with the systems version.
* Help: Wil respond with a hard-coded ‘help’ text.
* Identify: Will respond with identifying features of the user and the current conversation, such as user ID and platform in use.
* Ping: Will tell the user if a given IP address or endpoint is online and replying to pings.
* Traceroute: Will perform a traceroute to a given endpoint, and reply with individual messages to the user with each hop.
* SSH: Will open up a live shell session with a given server, asking for username, password, and port and then forwarding messages to and from the server.
* Logs: Will connect to a hard-coded server and allow the user to choose (with natural language) between several different logs, before tailing the chosen log to the user.
* Associate: Will associate this conversation with a hardcoded query, which operates on a hardcoded timer. If an ‘alert’ is triggered by this query, the user will be informed and the alert data will be sent to them.

It is believed that these intents are sufficient examples of what the framework can do and how it can be worked with.

### Intent Processing(IP) Module

The IP module manages all complex processing of information, and is called upon by several modules. At its core, the IP module receives data to the required function, processes information and then returns that information via callback to whichever module requested it. Not all intents require extended processing, but it makes sense to separate the majority of the processing that happens away from direct user responses.

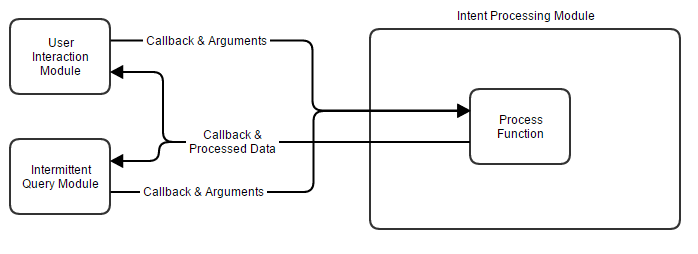


Figure 5 - Intent Processing Data Flow Example

Figure 5 shows an example of the flow of data from two modules to the Intent Processing module. In this case, either module can request a function from the IP module as long as it has the correct arguments and provides a callback, the IP module will then reply with the processed data.

Referring back to our list in the UI module planning, only the Ping, Traceroute, SSH and Associate intents will require processing in the IP module. Additionally, there will be a ‘helper’ file downloading function, which will download data from at least SSH servers for log retrieval. Each function will require specific data from the UI module, and will return specific processed data.

Ping can expect simply an address and a callback. It will send a request to the given address and call back whether the address is alive or not. This is our simple test to show that the framework is able to take a command, process it, and return information to the user. Traceroute, similarly, can expect only an address and a callback, returning the call for every single hop. This shows that the application has the ability to return multiple times to send the user better formatted data. Associate will expect three arguments: Some data referring to which query the user intends to associate the conversation with, the session variable which allows the function to save it against the query, and the callback in order to inform the user of the success or failure. Finally, the SSH function will be able to take the previously compiled connection data to connect to a server, and receive new data once a connection is established to forward messages from the user to the server, as well as reply with server messages through the callback. SSH will be a complex example of multi-stage conversation and will be a challenge to see if it is even possible with the planned configuration.

### Data Parsing Module

This module parses incoming data from servers into data more easily readable by other modules, and attempts to keep data consistent in the case of module replacement. The Parsing module will have functions added as required, and will primarily use Regular Expressions and replacement functions to make data more consistent. For example, the Server Query module may receive log data from a server containing line by line date and login time data, it could then send this data to the Data Parsing module to format this data into an object with a more Javascript friendly date system plus having the data predictably located in a variable as part of the object tree. However, this function would have to be hardcoded in as part of the Data Parsing module.

This module will be used when we deem necessary while developing, rather than being rigorously planned at this stage. An example data flow is shown in Figure 6, where various modules send unformatted data to the Data Parsing module, and the module returns back formatted data. This makes it significantly easier to maintain cleaner code when it comes to managing data and text by keeping all major text manipulation in one place.

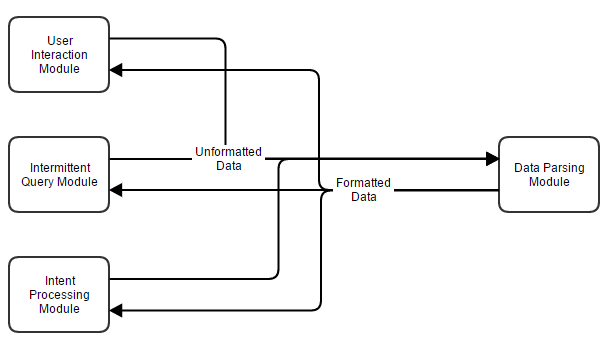


Figure 6 - Data Parsing Module example data flow

### Server Query Module

The Server Query module manages queries to and from servers. It will have a function for each major file transfer protocol, such as SCP and FTP, which will both accept connection information and file locations to be downloaded. These functions will regularly be called by the Intermittent Query Module in order to make a request to a server, and this data will parse through the Data Parsing module to enter a consistent format.

In our example intents, the only functions that actually do regular queries to server are the ‘log’ intent and the Intermittent Query Module, which contains a hardcoded list of queries the application does regularly. The Server Query module will be quite simple in practice, just being functions able to download files using data provided by other modules.

An example function for SCP would import an SCP downloading Node module, and expect an object containing the server address, file location and a local location where to save that file, and would call back to the function that called it with the downloaded file location. This function can then retrieve the data and send it to the Data Parsing module if required.

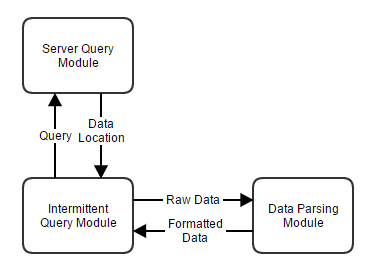


Figure 7 - Intermittent Query, Server Query & Data Parsing Interaction Example

As shown in Figure 7, the Intermittent Query module will sent a query to the Server Query module containing required data, such as credentials and address information, and the Server Query module will call back with the location of the data. The Intermittent Query module can then read that data and send it to the Data Parsing module if it needs to.

Separating Server Queries from the other module makes sense at this stage as it can be highly customised. It is likely that server queries will be written from scratch depending on differing systems and are unlikely to be shared between installations, unless a large number of generic queries can be created. It is important therefore that the Server Query module be very generic in its structured responses, so it can be easily replicated when creating more queries. Figure 8 contains a snippet of the proposed structure of data output from the module. Very generic, containing just the name of the query (for confirmation), the data requested and any other information just in case it is relevant.

{

**"name"**:"serverQuery1",

**"data"**:"values",

**"other"**:"OtherInfo"

}

Figure 8 - Server Query Example Output

### Intermittent Query Module

This module queries data from servers based on a timer, and is related primarily to the “monitor” mode of the program. Server queries will be stored as objects inside of an array, and hard coded into the program.

An example would be an intermittent query that checks, every minute, the ‘auth’ log from a Linux server. In this case, we would need to SCP to the server every minute and download the auth log, then parse it for information we cared about such as failed or successful logins, being careful to not include information of our own connection to get data, and then inform the user if it meets certain criteria. This also means the user must associate their conversation with the query using the ‘associate’ intent before receiving alerts.

{

name: 'LinuxAuthLogs',

server: 'linux', *//Defined in the servers file and looked up later*

type: 'logs',

application: 'system',

specific: 'auth’,

delay: 60000, *//How often between queries*

run: **function**(callback) {

*//Call the server query module to get the file*

serverQuery.scp(**this**.server, **this**.type, **this**.file, **function**(fileLocation) {

*//Call back with file location and read the file*

readfile(fileLocation, **function**(file) {

*//Call back with data and parse it*

**var** parsedFile = dataParse(file);

*//Inform the user using their saved session of the alert and its details*

informUser(parsedFile, userSavedSession);

})

});

}

}

Figure 9 - Example layout of an intermittent query

Figure 9 shows an example layout of an intermittent query in pseudocode, detailing all information related to the server aside from credentials, which is saved in the servers file shown later. It contains a function that can be called, which handles the gathering of data using other modules and then finally responding to the user. The name is used to associate users, where a user can simply type ‘associate LinuxAuthLogs’ to receive alerts for that query. The function itself will be called by a separate part of the module, which loops through queries on start-up and configures the intervals at which the queries will be run automatically based upon the definitions above.

### Stored Data

Data that may be required to be stored between sessions will be stored on disk in the form of JSON. Data that may be relevant to store include conversational sessions so users do not have to re-associate themselves with a query, and server data such as log history. If data needs to be stored, we can simply call JSON.stringify() on it and save it to disk, then load it to a variable on startup. This will be done as needed and is not a directly planned occurrence.

### Server File

The server file is simply a file containing objects of known servers, their logins and known paths to logs. These objects can be navigated by the Server Query module to obtain information required to retrieve logs and other data.

{

"linuxServer": {

address: 'linux.local',

protocol: 'scp',

port: 22,

username: 'root',

password: 'toor',

logs: {

system: {

auth: {

path: "/var/log/",

file: "auth.log"

}

}

}

}

}

Figure 10 - Server object example

Figure 10 shows an example of a simple server layout. A correctly configured Intermittent Query would be able to access the full path of the log file, as well as the address, protocol, port, and login by simply following the object. For example, *servers[this.server][this.type][this.application][this.specific].path* when referencing from the Intermittent Query module would give the module access to the path for that server.

### Setup File

The setup file will not be included in the final project paper as it contains private information such as passwords to servers, private keys for the LUIS application and lists of whitelisted IP addresses. This seemed necessary to address as the mysterious ‘setup’ variable is reference often in the code appendices but never makes an appearance itself. It is simply an object containing information not directly relevant to the code.

### LUIS

The choice of using the LUIS NLP service between others was a tough one at first, but it was decided that it in fact did not really matter which NLP service was chosen. As the project is deemed to be a framework rather than a fully defined software application, LUIS can theoretically be replaced with any one of many NLP services as long as the NLP module is tuned to output the same information.

However, LUIS ended up being the chosen service for testing due to the project lead having prior experience with it, and it seemed fairly fast to train and export learning data to a remote Azure server. The choice of Azure was easy, as it naturally connects to LUIS and Bot Builder, allowing for a much-simplified approach to the project.

# Implementation

This section will show all the trouble I had building my application, such as NLP training and SSH not acting like I expected at all.

## Forwarding Ports

The first step in producing the application was met with the issue of forwarding ports. Because Bot Framework expects you to have a HTTP endpoint for it to send information to, and one had not yet been set up on a production server, we have to open ports locally. This is a problem, as our test environment does not allow for port forwarding.

To fix this we used an application called Ngrok[5]. Ngrok allows us to expose local servers to the internet from behind a NAT when we are not able to port forward. Using this tunnel allowed us to continue our test deployment.

At this point, we have an internet-connected chat bot that we can test with the Bot Framework Emulator[6].

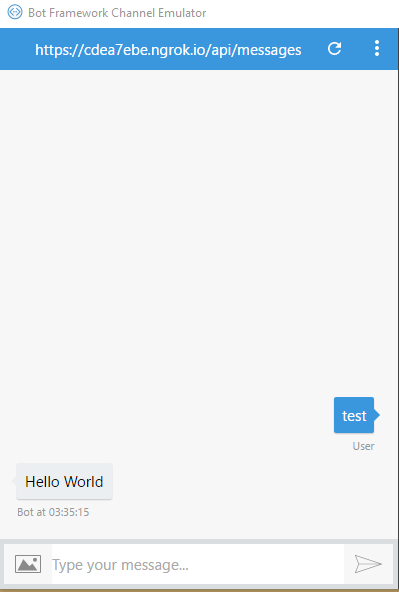


Figure 11 - Successful remote call and response

## IM Connection

Next we need to connect our application to an IM service. In this case Telegram, as it has an abundance of formatting styles and rich messaging capability that can be used to make data look better for the user, if required. This is as simple as registering a user on Telegram and then entering it into Microsoft’s API.

We now have access to our bots’ commands from within Telegram.

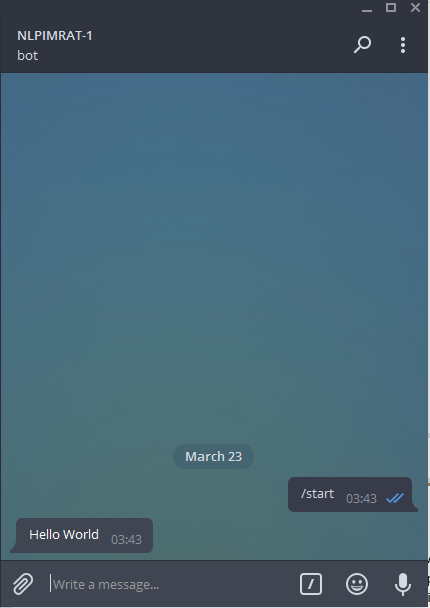


Figure 12 - Call and response from within Telegram

## Implementing Natural Language Processing

NLP must be added early, even prior to training being started. It is important that the module is at least configured so it can be expanded upon easily.

All programming for Luis.ai is done via their website, but when finished is hosted on an Azure endpoint. After setting up a basic NLP app and publishing it to Azure, we have access to the ‘intents’ that we configured. Intents are defined as “What the user intended to do based on what was said”, and data is formatted in a way that is understandable in code rather than in natural language. These intents allow us test our connections between Azure, BotBuilder, and Telegram.

In this case, some very simple training had to be added for Luis to understand the ‘ping’ command.

Training is a fairly simple process, but can become very intensive. To add a command first you must break the command down into its most bare components. In this case, the smallest possible rendition of the command is “ping {$address}”. But in natural language, it could be “Could you please ping the address {$address} for me?”. We must think up as many different forms of this as possible to train LUIS into understanding what it receives.

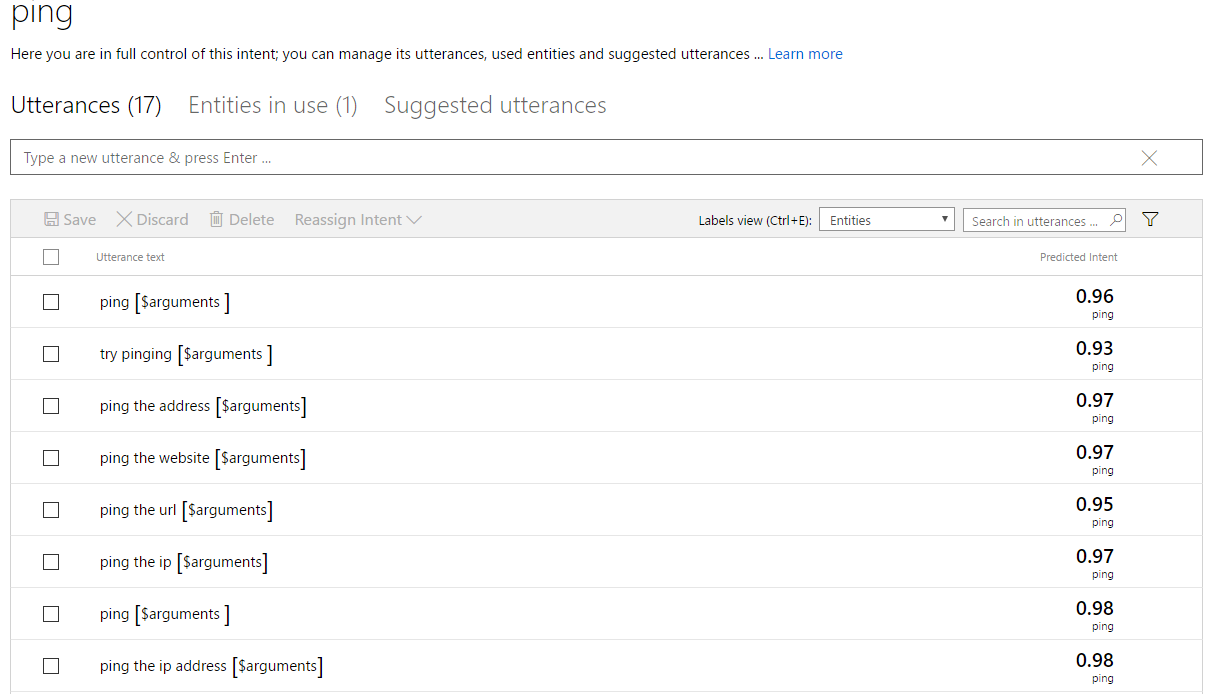


Figure 13 - Training the 'ping' intent

In Luis, we named our intent “ping”, and we can easily tell if a user’s intent was to ping by the response we receive from LUIS. For example, if it DID match ping, we can tell the user what arguments we managed to pull from their command. Failing that, we can give a default error message. At the same time, we can also program in the intents for querying the version number, and some templates for a help system.

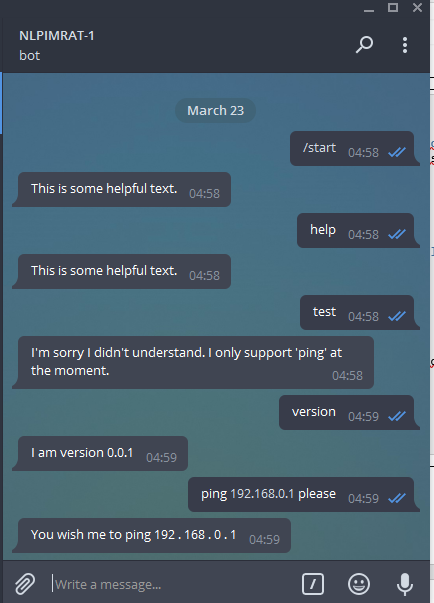


Figure 14 - Initial NLP training

After a successful conclusion to the NLP testing, point we started moving into a more modular and dynamic design. It is important that the NLP module be able to be replaced while leaving the intent processing intact, which is why the majority of the intent-processing capabilities of the module be generated by the UIM.

A simplified test version of the code can be seen in Figure 6 - User Interaction Module(UIM) example code. This code snippet shows a very simple example of the ‘ping’ command. We know that ‘ping’ is an intent that might be returned by the NLP module, so will respond with our formatted response of the data.

*//Grabs known intents from user interaction module, creates a template handler for each by looping through them*

**for** (**var** i = 0; i < knownIntents.length; i++) {

**var** match = knownIntents[i];

**var** intentGenerator = `

dialog.matches('*${*match*}*', [

function (session, args) {

console.log("Handling intent: *${*match*}*")

var arguments = [];

//Simplifies arguments for the intent handler

for (var i = 0; i < args.entities.length; i++) {

arguments.push({

name: args.entities[i].type,

value: args.entities[i].entity

});

}

//Sends intent plus data to the UIM for processing, then returns back with the

//text for the user

intentHandler['${match}'](arguments, function(response) {

session.send(response);

}); }

]); `;

eval(intentGenerator);

}

Figure 15 - NLP Module template generator

*//List of known intents*

*//This allows us to have a static list of expected data even in the case that the NLP module changes*

**var** knownIntents = ['ping'];

*//Handling of each intent, NLP module agnostic*

**var** intentHandler = {

ping: **function**(arg) {

console.log(arg);

**var** response = "You wish me to ping " + arg[0].value.replace(/\s/g, "");

**return** response;

}

};

Figure 16 - User Interaction Module(UIM) example code

In the case of the NLP module, handling this data is slightly more complex, but infinitely expandable. As it generates all of handling templates and rules from the UIM, it can be replaced easily with an NLP module for another service, instead of using LUIS. Figure 6 - NLP Module template generator shows this templating engine in action, by generating the endpoints for all the intents from LUIS and formatting the data to a generic format to be managed by the UIM.

## Implementing Basic Commands

Commands such as ‘ping’ and ‘traceroute’ are very simple in theory, but can be incredibly powerful to a user who is not in the network. Our Intent Process Module (IPM) manages the more advanced intents that require interacting with other modules.

### Ping

Instead of using the systems in-built ping command by piping user input straight to a console, which would make it ripe for command injection, we instead opted to use a powerful but small library simply named ‘ping’. It sanitises inputs and simply checks if a given IP is online or not. Using Node’s in-built DNS tools combined with *ping* allows us to quickly check if a given address is ping-able, as well as using our previously trained NLP service, we can set up a small alive-or-not host checker.

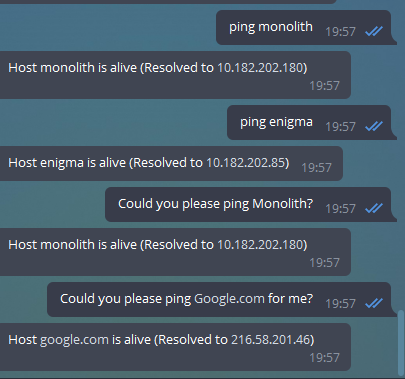


Figure 17 - Simple host PING example

### Traceroute

done

### SSH

## Advanced Server Querying

### Log Handling

### Automation

## Alerts System

## Securing the Software

This section will show how I secured the software from use by unauthorised parties

## Testing

This section will show testing against general use.

# References

[1] Ubuntu, “LinuxLogFiles - Community Help Wiki,” 2015. [Online]. Available: https://help.ubuntu.com/community/LinuxLogFiles. [Accessed: 23-Mar-2017].

[2] Microsoft, “Microsoft Bot Framework.” [Online]. Available: https://dev.botframework.com/. [Accessed: 23-Mar-2017].

[3] Restify, “API Guide | restify,” 2016. [Online]. Available: http://restify.com/. [Accessed: 14-Apr-2017].

[4] LUIS, “LUIS: Homepage,” 2017. [Online]. Available: https://www.luis.ai/home/index. [Accessed: 05-Mar-2017].

[5] A. Shreve, “ngrok - secure introspectable tunnels to localhost,” 2017. [Online]. Available: https://ngrok.com/. [Accessed: 23-Mar-2017].

[6] Microsoft, “Bot Framework Emulator | Documentation | Bot Framework,” 2017. [Online]. Available: https://docs.botframework.com/en-us/tools/bot-framework-emulator/. [Accessed: 23-Mar-2017].