Drexel Chatbot Project Plan

Hoa Vu <htv27@drexel.edu>
Tom Amon <tpa27@drexel.edu>
Daniel Fitzick <dwf35@drexel.edu>
Aaron Campbell <ajc382@drexel.edu>
Nanxi Zhang <nz66@drexel.edu>
Shishir Kharel<sk3432@drexel.edu>

Version: 3.0

Table of contents:

1. Revision History

Name(s)	Date	Comment(s)	Version
Everyone	10/09/16	Preliminary Project Plan	0.1
Tom, Hoa	10/18/16	Updated schedule	0.2
Hoa	10/19/16	Updated Statement of Work and Glossary	0.3
Everyone	10/30/16	Updated Project Background, Resource List, and Project Schedule	1.0
Everyone	1/17/2017	Updated Project Schedule and database resource	1.1
Everyone	1/18/2017	Updated Project Schedule and added demo as deliverable	2.0
Everyone	4/11/2017	Updated Project Schedule	3.0

2. Project Background

2.1 High-level description

This project is an AI chatbot (Drexel natural language query service) for Drexel. It will take questions about Drexel and provide an appropriate response. For example, when it receives the question "What time does the gym close today?", it will give a response “The gym closes at 10pm today”.

The system will be able to answer questions about the following topics:

- Drexel Facilities
- Drexel Staff
- Drexel Policies
- Food on Drexel’s campus
- Classroom locations
- Events

See the Customer Requirements Document for more information about features.

There will be three main AI components to the system:

- Information retrieval agent - capable of finding information through structured or unstructured websites, and storing information in a database
- Natural language parsing agent - capable of taking a natural language question and converting it to a machine understandable query
- Response generation agent - capable of using the information in the database to generate a natural language answer to the question

2.2 Research

Because this project requires the use of AI, it will be research focused. Many different AI algorithms can be used to accomplish the same goal, so some experimentation is required to find what will work the best for this exact situation. Professor Marcello Balduccini will guide the team to an understanding of the different learning algorithms and the final design of the chatbot.

3. Statement of Work

3.1 Web service API

An API will be developed so that other developers can leverage it to create their own applications. The API will be a webservice, where a user sends a question to a web address, and the web address responds with the answer to that question. This is the main deliverable of the project.

3.2 Example front-end applications

Two simple front end applications will be developed to demonstrate the web api's use: a web site and an Android application, developed by Aaron. Time permitting, a third test application: SMS, will also be added, developed by Daniel.

3.3 Demo

There will be a demo presented. This demo will be a webpage where a user can insert a string and press a button to send the string to the Drexel Chatbot API, which will be running on a server. The API will then reply in JSON, which the webpage will translate to English and show it to the user. The system will only be trained to handle questions relating to Drexel's faculty.

3.4 Other Deliverables

- Project plan (Lead by Hoa)
- Requirements document (Lead by Daniel)
- Design document (Lead by Thomas)
- Code (Lead by Aaron)
- Bug report (Everyone)
- Test plan (Lead by Nanxi), including unit testing, as well as a specific number of questions asked from each category to generate statistics regarding correct responses.

4. Resource List

- Web server: Cloud At Cost
- Version control system: Github
- Issues and bug tracking: Github Issues
- Project management tool: Microsoft Project
- Programming language: Python
- Static analysis tool: Pylint
- Code coverage: Coverage.py
- Neural network library for python: Keras
- Natural Language processing library for python: NLTK(Natural Language Toolkit)
- Database: Stardog (semantic graph database)

5. Assumptions

5.1 Language

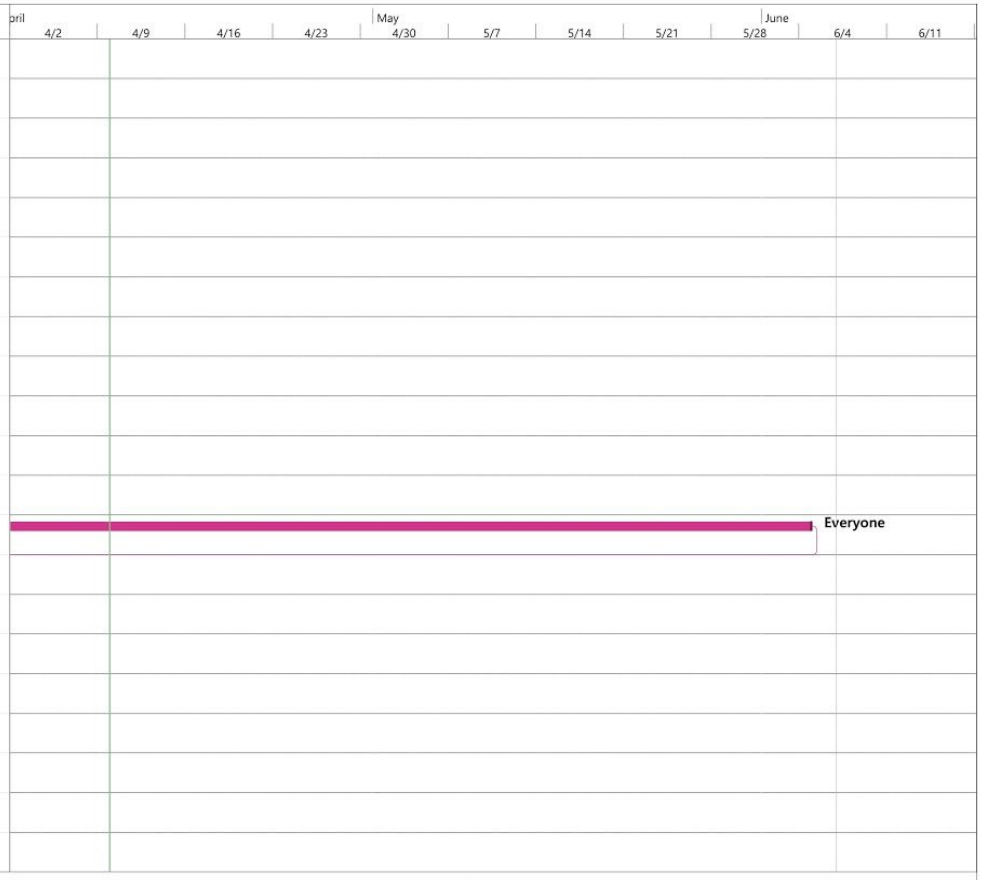
The user will be able to read and write in English, as this is the only language we will develop for.

5.2 Limited Scope

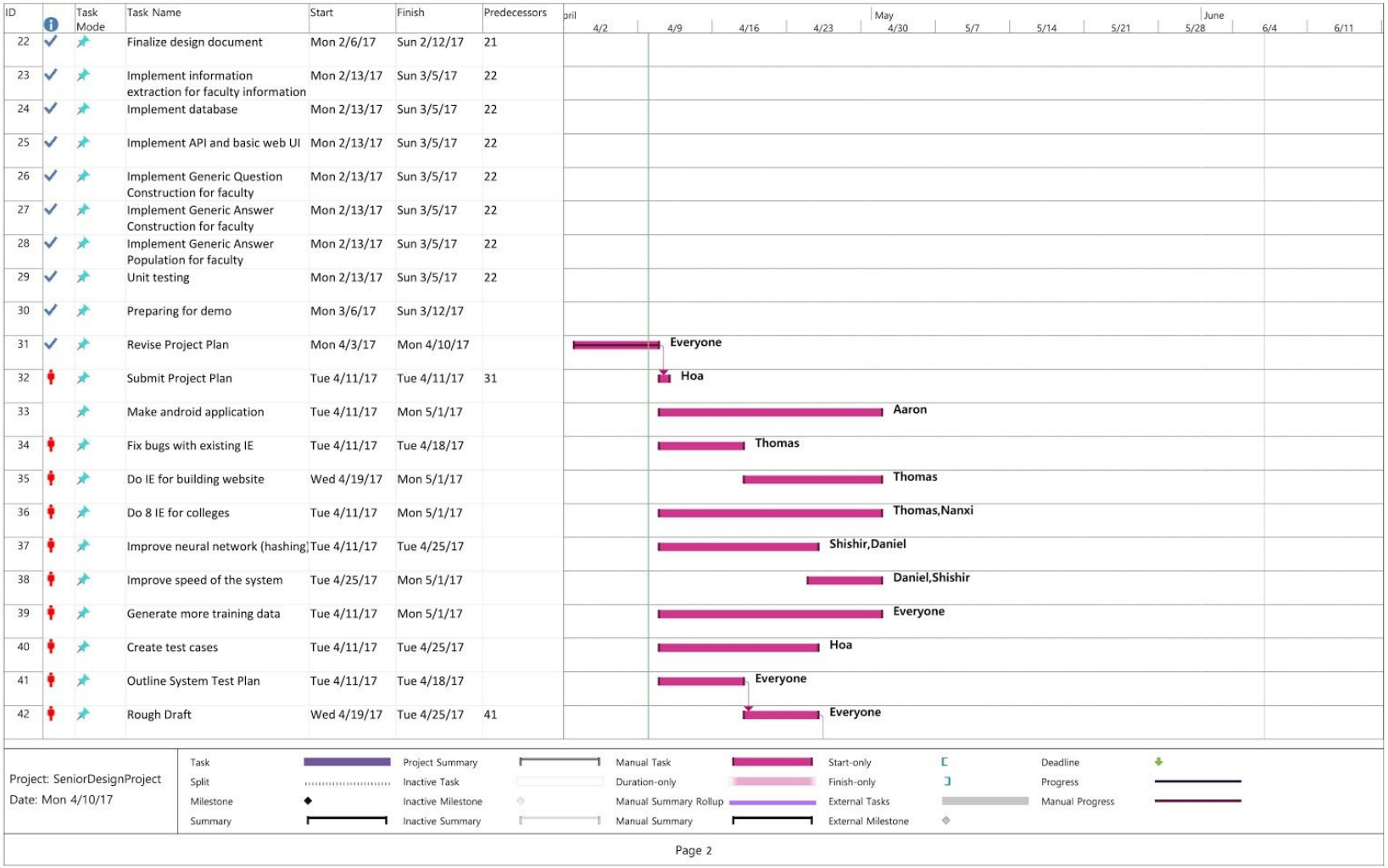
The user will only ask questions related to Drexel policies, services, etc.

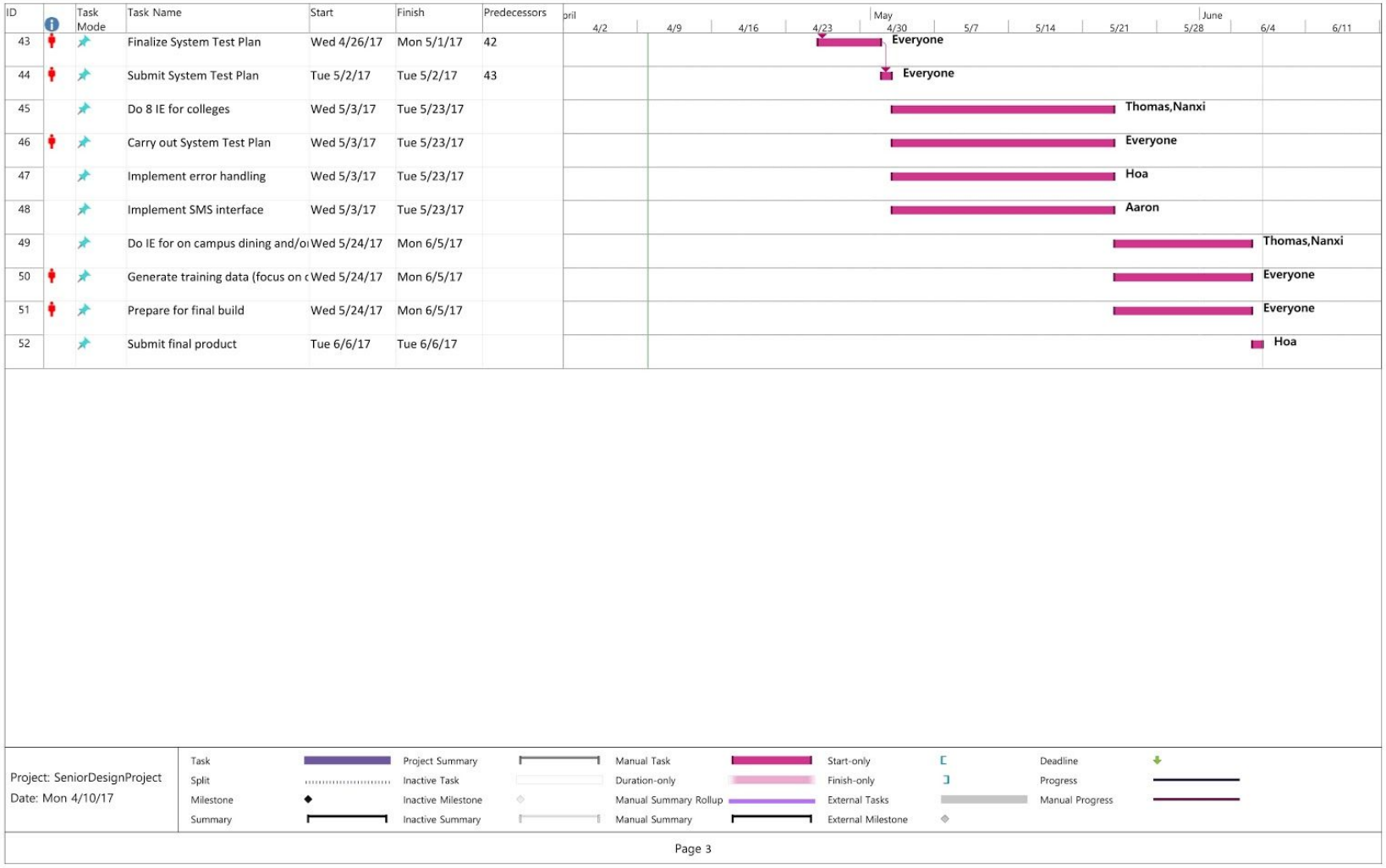
6. Project Schedule

ID	Task Mode	Task Name	Start	Finish	Predecessors
1		Prel. Project Plan	Tue 10/4/16	Sun 10/9/16	
2		Prel. Customer Requirements	Mon 10/10/16	Sun 10/16/16	
3		Meet Stakeholder	Thu 10/20/16	Thu 10/20/16	
4		Customer Requirements (baseline	Mon 10/17/16	Sun 10/23/16	
5		Second Draft Project Plan	Mon 10/10/16	Sun 10/23/16	1
6		Project Plan (baselined)	Sun 10/23/16	Sun 10/30/16	5
7		Research database	Mon 10/24/16	Sun 10/30/16	
8		Set up database and website	Mon 10/31/16	Sun 11/13/16	
9		Research Requirement Spec. & Create Prototype	Mon 10/24/16	Wed 11/16/16	
10		Prel. Requirement Spec.	Mon 11/14/16	Sun 11/20/16	
11		Requirement Review	Wed 11/23/16	Tue 11/29/16	
12		Requirement Spec. (baselined)	Tue 11/29/16	Sun 12/4/16	11
13		Weekly meetings online	Mon 1/9/17	Sun 6/4/17	
14		Revise project plan	Tue 1/10/17	Sun 1/15/17	13
15		Finalize project plan	Mon 1/16/17	Sun 1/22/17	14
16		Make generic answer formats and list of generic	Mon 1/16/17	Sun 1/29/17	
17		Learn how to use Stardogs	Mon 1/16/17	Sun 1/29/17	
18		Set up work environment + continue tryout	Mon 1/16/17	Sun 1/29/17	
19		Gather training data for Neural Network testing	Mon 1/16/17	Sun 2/5/17	
20		Weekly meeting on campus	Wed 1/18/17	Wed 2/8/17	
21		Rough draft for design document	Mon 1/16/17	Sun 2/5/17	14



Page 1


ID
Task Mode
Task Name
Start
Finish
Predecessors
April
May
June
4/2
4/9
4/16
4/23
4/30
5/7
5/14
5/21
5/28
6/4
6/11
22 Finalize design document Mon 2/6/17 Sun 2/12/17 21
23 Implement information extraction for faculty information Mon 2/13/17 Sun 3/5/17 22
24 Implement database Mon 2/13/17 Sun 3/5/17 22
25 Implement API and basic web UI Mon 2/13/17 Sun 3/5/17 22
26 Implement Generic Question Construction for faculty Mon 2/13/17 Sun 3/5/17 22
27 Implement Generic Answer Construction for faculty Mon 2/13/17 Sun 3/5/17 22
28 Implement Generic Answer Population for faculty Mon 2/13/17 Sun 3/5/17 22
29 Unit testing Mon 2/13/17 Sun 3/5/17 22
30 Preparing for demo Mon 3/6/17 Sun 3/12/17
31 Revise Project Plan Mon 4/3/17 Mon 4/10/17 Everyone
32 Submit Project Plan Tue 4/11/17 Tue 4/11/17 31 Hoa
33 Make android application Tue 4/11/17 Mon 5/1/17 Aaron
34 Fix bugs with existing IE Tue 4/11/17 Tue 4/18/17 Thomas
35 Do IE for building website Wed 4/19/17 Mon 5/1/17 Thomas
36 Do 8 IE for colleges Tue 4/11/17 Mon 5/1/17 Thomas,Nanxi
37 Improve neural network (hashing) Tue 4/11/17 Tue 4/25/17 Shishir,Daniel
38 Improve speed of the system Tue 4/25/17 Mon 5/1/17 Daniel,Shishir
39 Generate more training data Tue 4/11/17 Mon 5/1/17 Everyone
40 Create test cases Tue 4/11/17 Tue 4/25/17 Hoa
41 Outline System Test Plan Tue 4/11/17 Tue 4/18/17 Everyone
42 Rough Draft Wed 4/19/17 Tue 4/25/17 41 Everyone
Project: SeniorDesignProject
Date: Mon 4/10/17
Task
Split
Milestone
Summary
Project Summary
Inactive Task
Inactive Milestone
Inactive Summary
Manual Task
Duration-only
Manual Summary Rollup
Manual Summary
Start-only
Finish-only
External Tasks
External Milestone
Deadline
Progress
Manual Progress
Page 2



ID	Task Mode	Task Name	Start	Finish	Predecessors
43		Finalize System Test Plan	Wed 4/26/17	Mon 5/1/17	42
44		Submit System Test Plan	Tue 5/2/17	Tue 5/2/17	43
45		Do 8 IE for colleges	Wed 5/3/17	Tue 5/23/17	
46		Carry out System Test Plan	Wed 5/3/17	Tue 5/23/17	
47		Implement error handling	Wed 5/3/17	Tue 5/23/17	
48		Implement SMS interface	Wed 5/3/17	Tue 5/23/17	
49		Do IE for on campus dining and/o	Wed 5/24/17	Mon 6/5/17	
50		Generate training data (focus on c	Wed 5/24/17	Mon 6/5/17	
51		Prepare for final build	Wed 5/24/17	Mon 6/5/17	
52		Submit final product	Tue 6/6/17	Tue 6/6/17	

Project: SeniorDesignProject
Date: Mon 4/10/17

Task
Split
Milestone
Summary
Project Summary
Inactive Task
Inactive Milestone
Inactive Summary
Manual Task
Duration-only
Manual Summary Rollup
Manual Summary
Start-only
Finish-only
External Tasks
External Milestone
Deadline
Progress
Manual Progress

Page 3

7. Risks

7.1 Research

At this point there is not a designated algorithm to implement the chatbot model. Research on this topic will be conducted. Testing on different models will also be done through prototyping.

7.2 Data Acquisition

Drexel website data may not be structured in such a way as to be conducive to automated collection and processing. There will be a component developed to attempt to collect this data. If that is unsuccessful, the data will have to either be retrieved manually or through some other structured source.

7.3 Accuracy

The system will not be 100% accurate due to the nature of AI. A goal will be to maximize accuracy.

8. Glossary

- Chatbot: An interface, usually text based, specializing in the mimicry of natural language conversation. AKA “artificial conversational entity”
- SMS: Short Message Service, a text messaging service of telephone.