

# A Conversational Ahatbot Architecture for eHealth Systems

Niccoló Terreri

August 5, 2016

# Contents

# Chapter 1

## Introduction

### 1.1 PEACH

#### 1.1.1 The chatbot application team

The search engine

The restful server



# Chapter 2

## Background Research

### 2.1 The electronic Health Needs Assessment questionnaire

#### 2.1.1 Macmillan Cancer Support

#### 2.1.2 The Concerns Checklist

### 2.2 Patient Data for Research in the UK

#### 2.2.1 Data Protection Act 1998

#### 2.2.2 Freedom of Information Act 2000

### 2.3 Natural Language Processing

#### 2.3.1 Sentiment analysis

#### 2.3.2 Text categorization

#### 2.3.3 Open Source NLP libraries

Natural Language Tool-Kit

Stanford CoreNLP

### 2.4 Categorizing text with the NLTK

#### 2.4.1 Supervised learning

#### 2.4.2 Development, feature extraction and evaluation

#### 2.4.3 Uncertainty bounds<sup>5</sup>

#### 2.4.4 Expected probability

#### 2.4.5 The limits of the NLTK and CoreNLP

#### 2.4.6 Conversations and sequence classification

# Chapter 3

## Requirements Gathering

### 3.1 The Conversational UI

#### 3.1.1 A hybrid model

#### 3.1.2 The conversation

#### 3.1.3 Topic management

#### 3.1.4 Report generation

### 3.2 Requirements Listing



# Chapter 4

## System Design and Implementation

### 4.1 Software Architecture

#### 4.1.1 The Internet is an IO device

#### 4.1.2 Principles of Software Design

#### 4.1.3 Deployability across chat clients

### 4.2 The internal structure

#### 4.2.1 Preprocessing

Tokenization

Stopwords

Stemming

Categorization

#### 4.2.2 User and topic management

#### 4.2.3 The RiveScript interpreter

#### 4.2.4 Postprocessing

Decorating the response

#### 4.2.5 Integration with the external search engine



# Chapter 5

## System Testing and Evaluation

### 5.1 Test Driven Development

Unit tests

Integration tests

### 5.2 Evaluating a categorizer

K-fold evaluation

Evaluating a sequence classifier

# Chapter 6

## Conclusion

### 6.1 Project Results

### 6.2 Team Retrospective

### 6.3 Future Work

# Appendix A

## System Manual

# Appendix B

## User Manual

# Appendix C

## Tests Listing

## Appendix D

### Categorizer Evaluation

# Appendix E

## Code Listing