

# Improving the Chatbot Fallback Experience

With a Content-Based Recommender System



# Introduktion av företaget [A]

— — —

- Globalt distribuerat företag
- Central Time (CT) 09:00 - 17:00
- <https://www.simplea.com>

*“[A] is a Content Intelligence Service that partners with leading global enterprises. [A] orchestrates content intelligence systems that unify people, processes, and technology for omnichannel publishing and real-time personalised customer experiences at scale.”*

*“Almost always [A] works under NDA. Often [A] handles highly-secured, regulated financial or medical data. Sometimes [A] works invisibly alongside agencies.”*


# Så här har jag jobbat


— — —


- 3 veckovisa, regelbundna möten i Zoom
- Kommunicerat genom Slack
- Möten och arbetstider bokas i Google calendar
- Tillgång till personlig Gmail
- Google Drive används för att hantera och dela dokument
- Lucidchart för att samarbeta visuella diagram
- Jira; Backlog, Analysis, Ready for Development, In Progress, Ready for Deployment, In Review, Done.
- Timetracking i Jira med Tempo

# Bakgrund till projektet, del 1

- Chatbot-tjänst
- Dialogflow
- Intent
- Natural Language Understanding, NLU

Training phrases 

Search training phrases 


” Add user expression 


” tell me about chatbot technology

” what is a chatbot



” I want to know more about chatbots

” What are chatbots good for?

Action and parameters 

Responses 

DEFAULT HANGOUTS +

Text response  

1 Quite simply, chatbots are the new FAQ. Every second, 40,000 search queries are made online worldwide. That's a lot of people looking for a lot of answers. Employing a chatbot lets you answer many of those inquiries.

2 Enter a text response variant

# Bakgrund till projektet, del 2

- Problemmotivering:  
Fallback intent  
Kundscenario / use case  
Chatbot log  
Jira Backlog

Conversation

4 interactions

Duration: 1 min

USER

Hi, how's it going?

AGENT

I'm doing very well. Thanks!

USER

Great. Please tell me about content engineering.

AGENT

Sure! Content engineering is the practice of organizing the shape, structure, and application of content. It bridges the gaps between strategy and development, and when applied; it results in sophisticated content that is designed for reuse.

USER

Okay, and what about chatbots? What are they?

AGENT

Quite simply, chatbots are the new FAQ. Every second, 40,000 search queries are made online worldwide. That's a lot of people looking for a lot of answers. Employing a chatbot lets you answer many of those inquiries.

USER

Interesting. How do they fit together, chatbots and content engineering? I mean, how can I use a chatbot in my content engineering strategy?

AGENT

No matched intent

# Bakgrund till projektet, del 3

- Recommender systems
- Huvudmålet:  
Recommender system  
ger stöd åt chatbot  
vid fallback intent  
så att chatbot agenten  
kan bidra med någon form  
av nytta till användaren

## Customers Who Bought This Item Also Bought



The screenshot displays a horizontal list of six recommended books. Each item includes a book cover, title, author, star rating, number of reviews, and price. The first item, 'Data Science from Scratch', is highlighted with a yellow background. Navigation arrows are visible on the left and right sides of the list.

Book Title	Author	Rating	Reviews	Price
Data Science from Scratch: First Principles with Python	Joel Grus	★★★★☆	54	\$33.99 ✓Prime
Python for Data Analysis: Data Wrangling with Pandas, NumPy, and...	Wes McKinney	★★★★☆	118	\$27.68 ✓Prime
Data Science for Business: What You Need to Know about Data Mining and...	Foster Provost	★★★★☆	135	\$37.99 ✓Prime
Reproducible Research with R and R Studio, Second Edition...	Christopher Gandrud	★★★★☆	3	\$51.97 ✓Prime
An Introduction to Statistical Learning: with Applications in R...	Gareth James	★★★★☆	105	\$68.35 ✓Prime
Data Smart: Using Data Science to Transform Information into Insight	John W. Foreman	★★★★☆	99	\$28.16 ✓Prime

Page 1 of 15

# Bakgrund till projektet, sista delen

— — —

- Projektmål: kontrollera om detta är en möjlighet som de kan jobba vidare på
- Önskat resultat →  
Länk till artikel med starkast relation till ämnet/ämnena som användaren efterfrågade + En paragraph som nämner dessa ämnen från artikeltexten

7 interactions  
Duration: 3 min

USER

Hi there, how's your day going?

AGENT

Doing great, thanks.

USER

Glad to hear it. Could you let me know about content engineering. What is that?

AGENT

Sure! Content engineering is the practice of organizing the shape, structure, and application of content. It bridges the gaps between strategy and development, and when applied; it results in sophisticated content that is designed for reuse.

USER

Thanks for the answer. And what are chatbots?

AGENT

It's my pleasure to help.

USER

What are chatbots?

AGENT

Quite simply, chatbots are the new FAQ. Every second, 40,000 search queries are made online worldwide. That's a lot of people looking for a lot of answers. Employing a chatbot lets you answer many of those inquiries.

USER

Sounds interesting. What are the ways I could use a chatbot in my content engineering strategy?

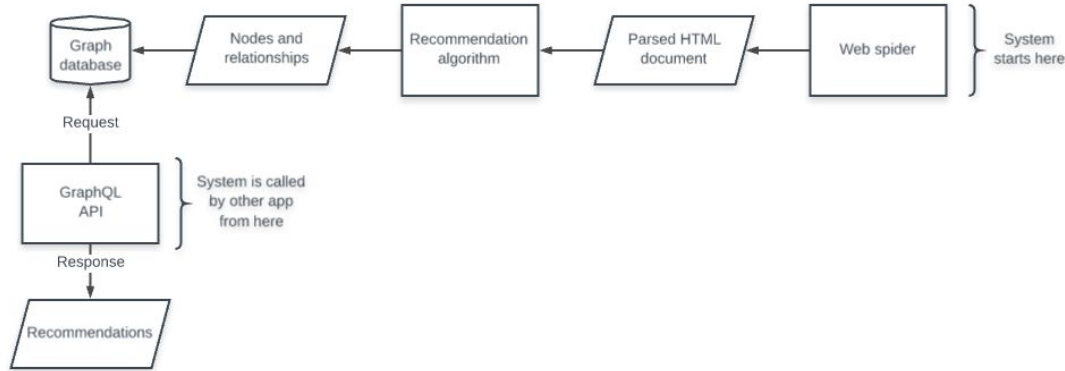
AGENT

I'm sorry, I don't fully understand what you mean but I did try to find you an article that might explain this topic to you. Take a look at this link: <https://simplea.com/Articles/AI-Marketing-Chatbots-and-Your-CMS> It seems to be talking about the topics you mentioned in this paragraph: "Download the [A] Guide to Unlocking the Chatbot Inside Your CMS — a two-part whitepaper on single-sourcing content for new interactive channels. Part 1 explains how to deliver content for chatbots and future on-demand channels using the content management system (CMS) you already maintain. Part 2 walks you step-by-step through the process of syncing, linking, marrying and integrating chatbot content with your CMS."

# Projektöversikt, del 1

— — —

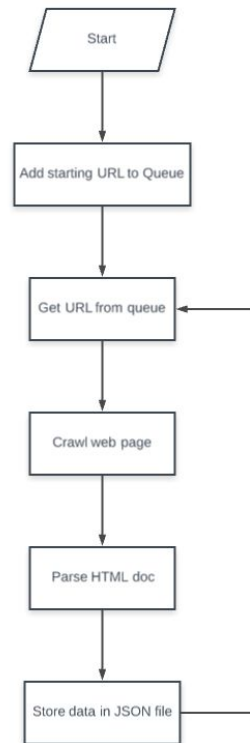
- Produktkrav: web spider (crawling och scraping), filtreringsalgorithm, databas, API



- Teknisk lösning: DotnetSpider, Bag-of-Words model, Neo4j graph database, GraphQL Query Language



\_\_\_\_\_



```

1 namespace Abot.Spider.Spiders
2 {
3     public class SimpleaSpider : DotnetSpider.Spider
4     {
5         protected override void Initialize()
6         {
7             NewGuidId();
8             Scheduler = new QueueDistinctBfsScheduler();
9             DownloaderSettings.Type = DownloaderType.HttpClient;
10            AddDataFlow(new SimpleaDataParser()).AddDataFlow(new JsonfileStorage());
11            AddRequests("https://simplea.com/Articles/facing-content-supply-chain-problems");
12        }
13
14        class SimpleaDataParser : DataParser
15        {
16            public SimpleaDataParser()
17            {
18                CanParse = DataParserHelper.CanParseByRegex("simplea\\.com/Articles");
19                QueryFollowRequests = DataParserHelper.QueryFollowRequestsByXPath('.');
20            }
21
22            protected override Task<DataFlowResult> Parse(DataFlowContext context)
23            {
24                if (context.Response != null)
25                {
26                    context.AddItem("URL", context.Response.Request.Url);
27                    context.AddItem("Title", context.GetSelectable().XPath("//title").GetValue());
28                    context.AddItem("Keywords", context.GetSelectable().XPath("//meta[@name='keywords']/@content").GetValue());
29                    context.AddItem("Summary", context.GetSelectable().XPath("//meta[@name='description']/@content").GetValue());
30                    var pTags = context.GetSelectable().XPath("//p").Nodes();
31                    int nr = 1;
32                    foreach (var p in pTags)
33                    {
34                        context.AddItem("Paragraph" + nr, p.GetValue());
35                        nr++;
36                    }
37                }
38                return Task.FromResult(DataFlowResult.Success);
39            }
40        }
41    }
42 }

```

# Projektöversikt, del 3

- BoW model hanterar den nedladdade informationen
- Normalisera texten
- “Ett ord är av stor betydelse om det nämns många gånger”
- Word count & frequency

“order”, [1, 6]

“give”, [1, 6]

“chatbots”, [2, 13]

“content”, [1, 6]

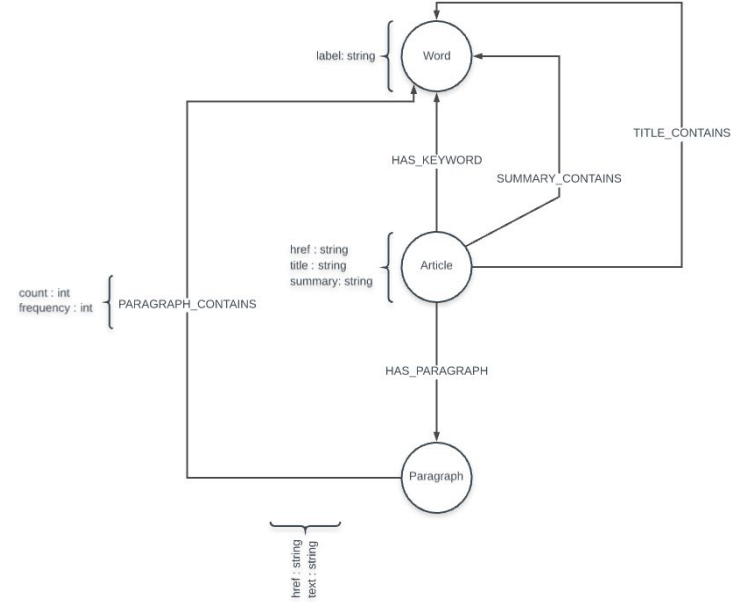
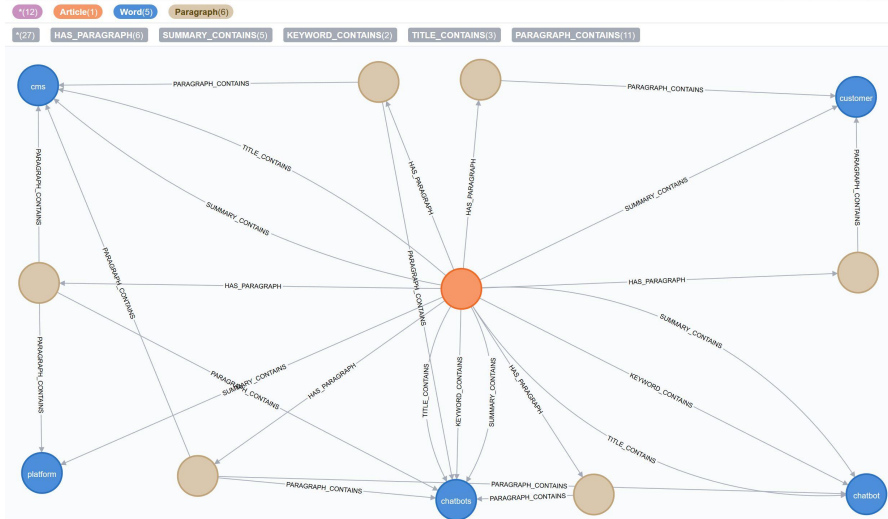
“understand”, [1, 6]

“work”, [1, 6]

```
foreach (Paragraph paragraph in article.Paragraphs)
{
    // 1. Get full article text by adding all paragraphs together
    fullArticleText += paragraph.Text + "\n";
    // 2. Split each paragraph into single words
    List<string> words = paragraph.Text.Split(" ").ToList();
    string stopwords = File.ReadAllText(_projectPath + "stopwords.txt");
    foreach (string word in words.ToList())
    {
        // 3. Remove whitespace and characters and convert to lower case for each word
        string normalizedWord = Regex.Replace(word.Trim().ToLower(), @"[^A-Za-z]+", "");
        words.Remove(word);
        words.Add(normalizedWord);
        // 4. Remove stopwords
        if (stopwords.Contains(normalizedWord))
        {
            words.Remove(normalizedWord);
        }
    }
}
```

# Projektöversikt, del 4

- Nodes (entities) och edges (relationships)
- Neo4j Driver for .NET
- Neo4j browser



# Projektöversikt, del 5

— — —

- GraphQL for .NET
- Get → Query
- Create, Update, Delete → Mutation

```
{  
  Word (label: "chatbots") {  
    Article (titleContains: true) {  
      url  
      Paragraph (score: 5) {  
        text  
      }  
    }  
  }  
}
```

# Avslutande tankar

— — —

- Tekniska krav gick att uppfylla
  - Absolut en idé som går att bygga vidare på
  - Production ready? Nej. Behövs absolut en mer avancerad algorithm alternativt ML model + tester och möjlighet att “fånga upp fel” saknas  $\Rightarrow$  svårt att scale.
- 
- Känts som ett otroligt omfattande arbete, svårt att hålla koll på allt...
  - Häftigt att man kan jobba tillsammans och befinna sig olika kontinenter, lite svårt att känna av “team work” känsla?

Tack!