

20-q

an online guessing game

What is 20-q?

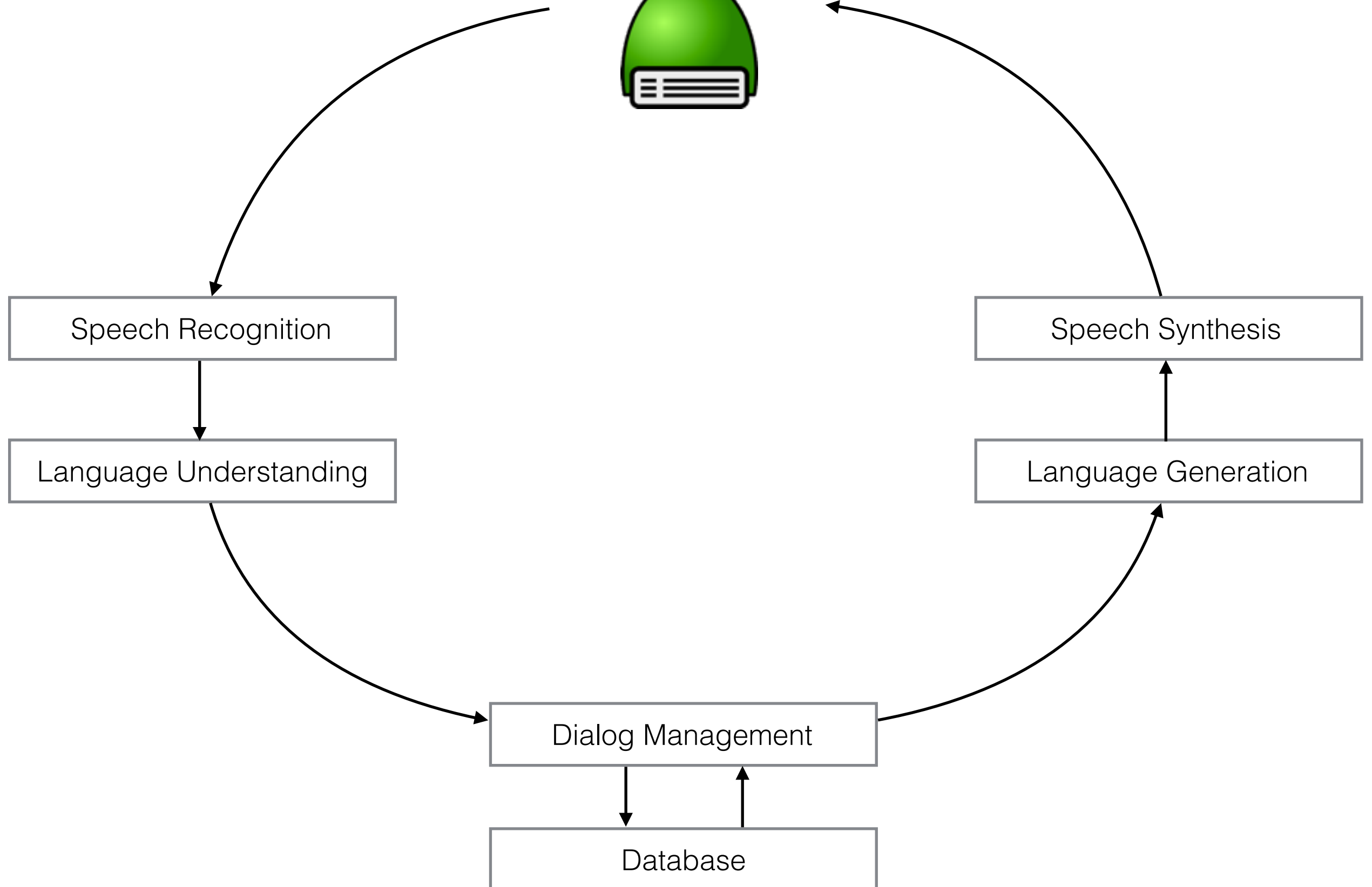
- 20 question game is a guessing game. One player (answerer) choose a subject and another player (questioner) asks at most 20 yes-no questions. If questioner find the answer, the questioner wins, otherwise the answerer wins.
- 20-q is an online 20 question game, which computer plays the questioner and human plays the answerer. With this limitation that subject should be a well known person.

Why?

it is important :

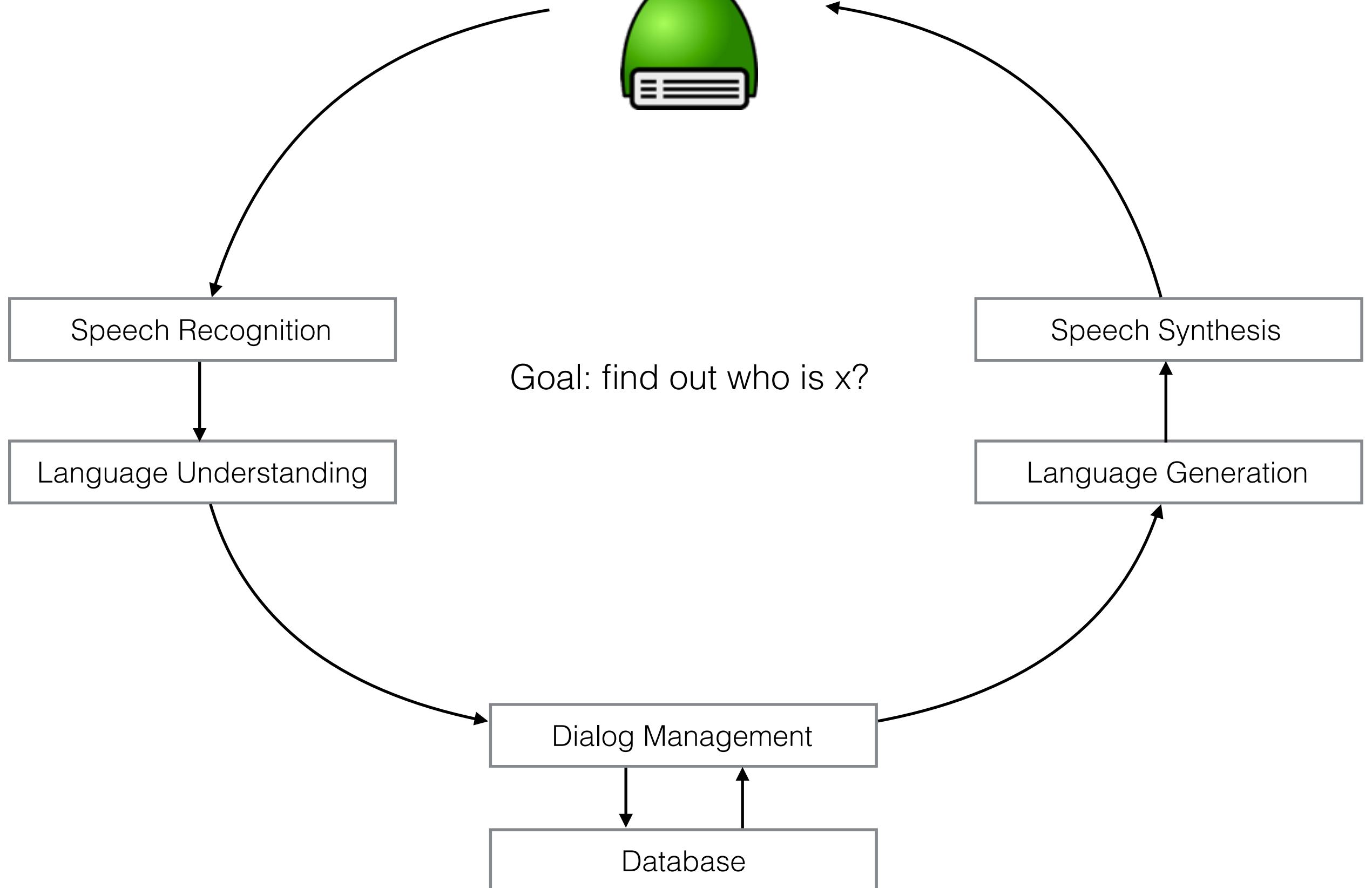
- As a search problem
- As a dialogue in natural language between human and computer

the dialogue system





Goal: find out who is x?





Goal: find out who is x?

$X == A ?$

Speech Recognition

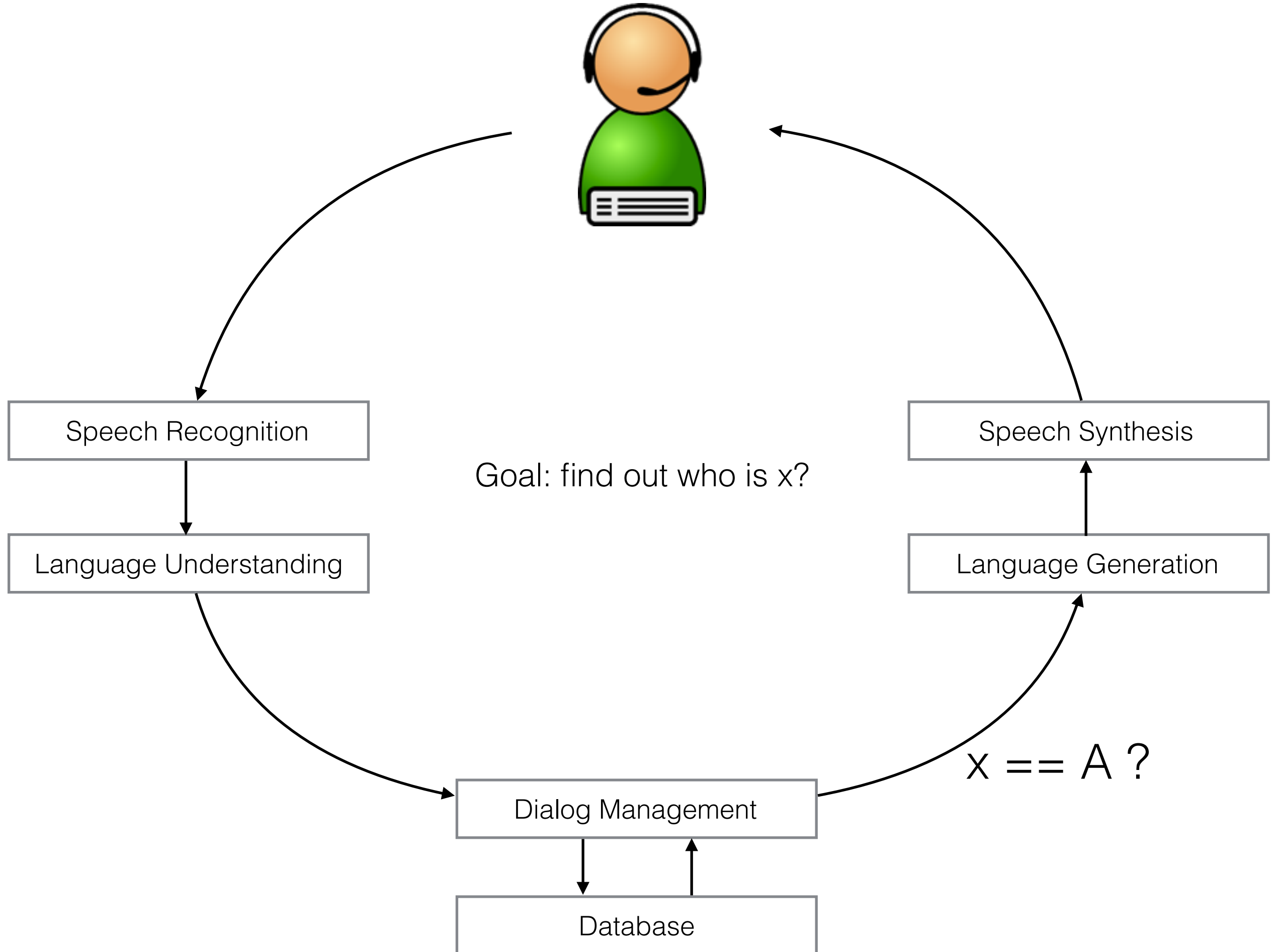
Language Understanding

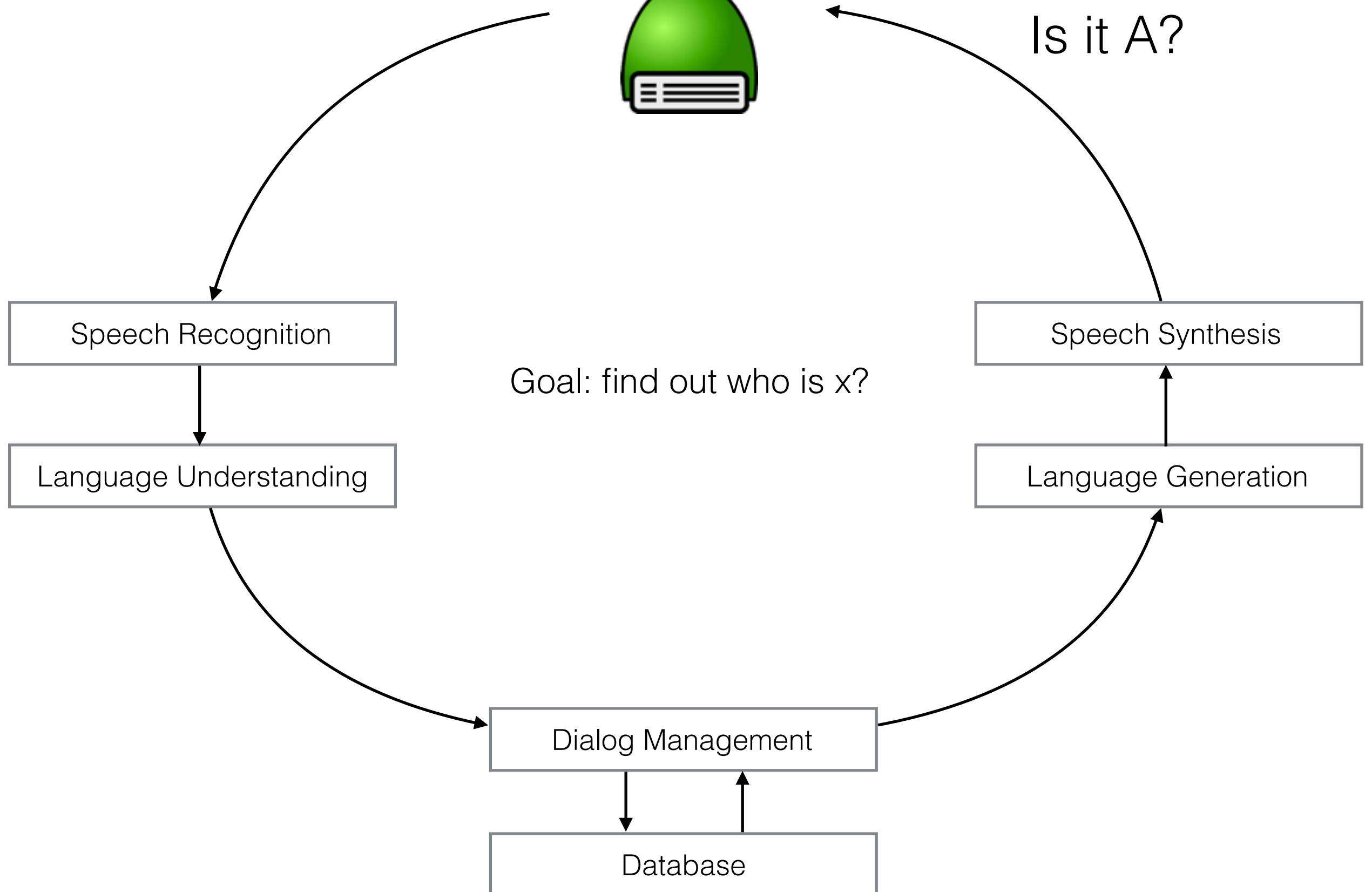
Dialog Management

Database

Speech Synthesis

Language Generation







No it is not A

Goal: find out who is x?

Speech Recognition

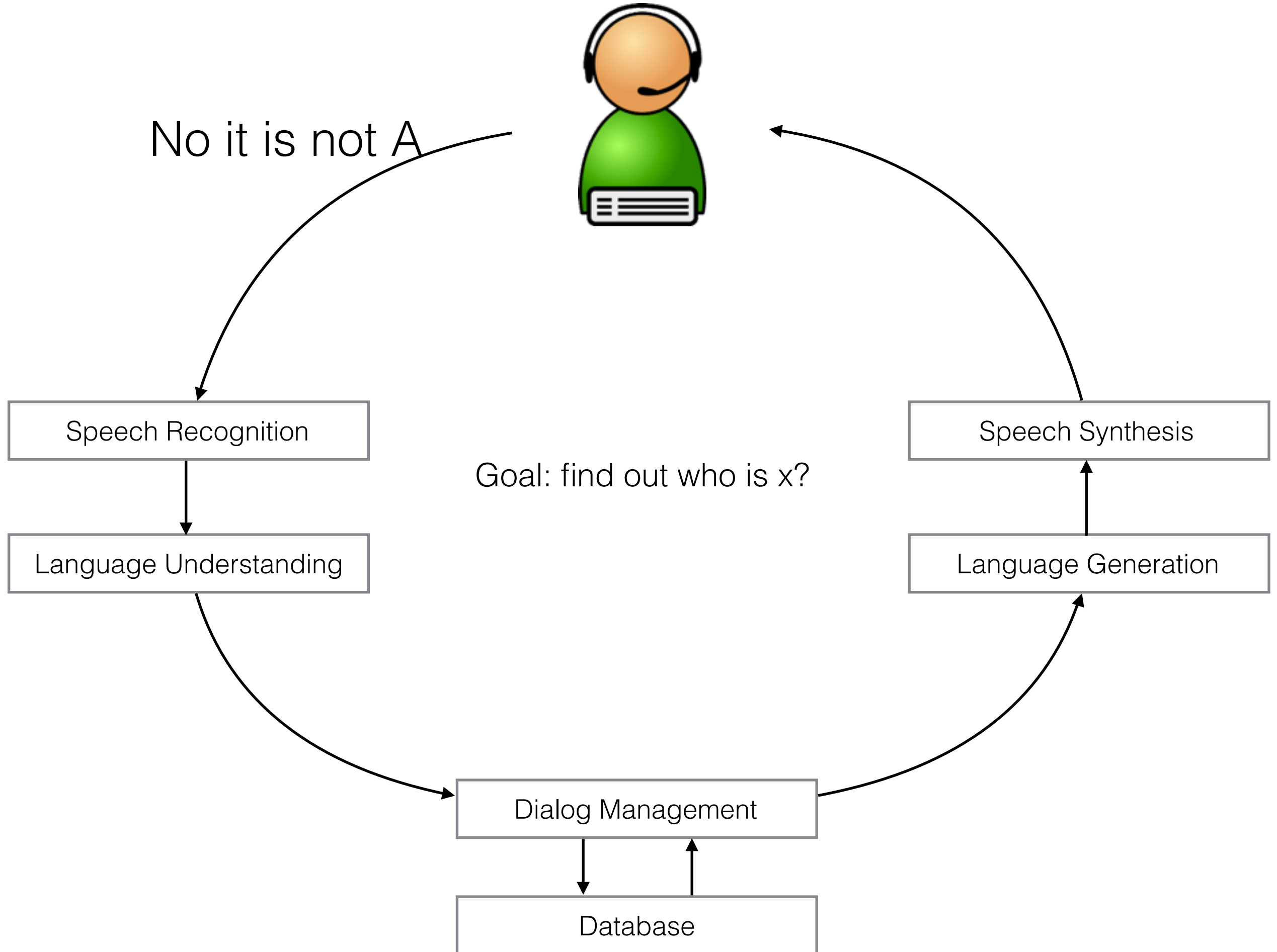
Language Understanding

Dialog Management

Database

Speech Synthesis

Language Generation





Goal: find out who is x?

Speech Recognition

Language Understanding

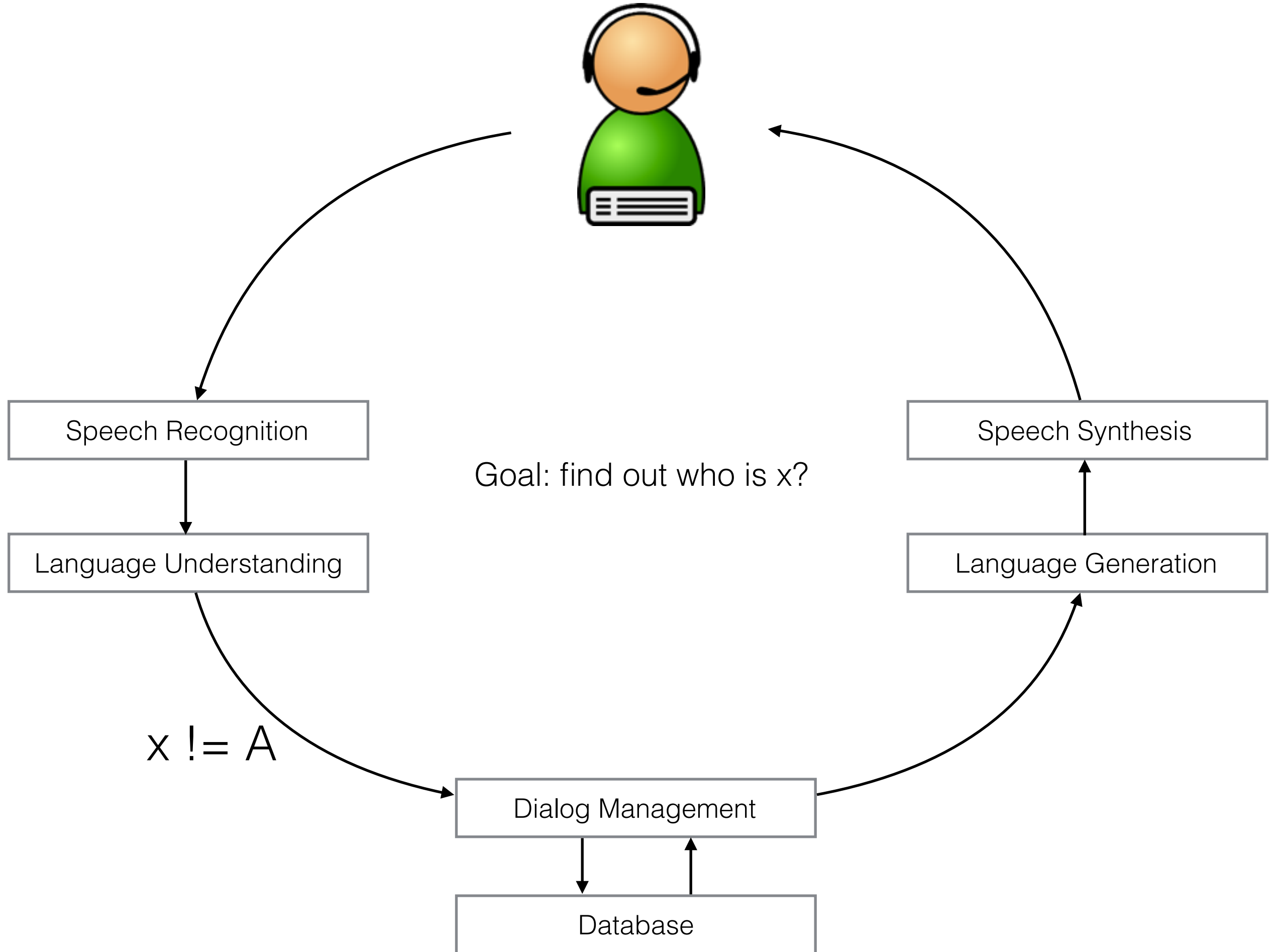
Speech Synthesis

Language Generation

$x \neq A$

Dialog Management

Database





Goal: find out who is x?

Speech Recognition

Language Understanding

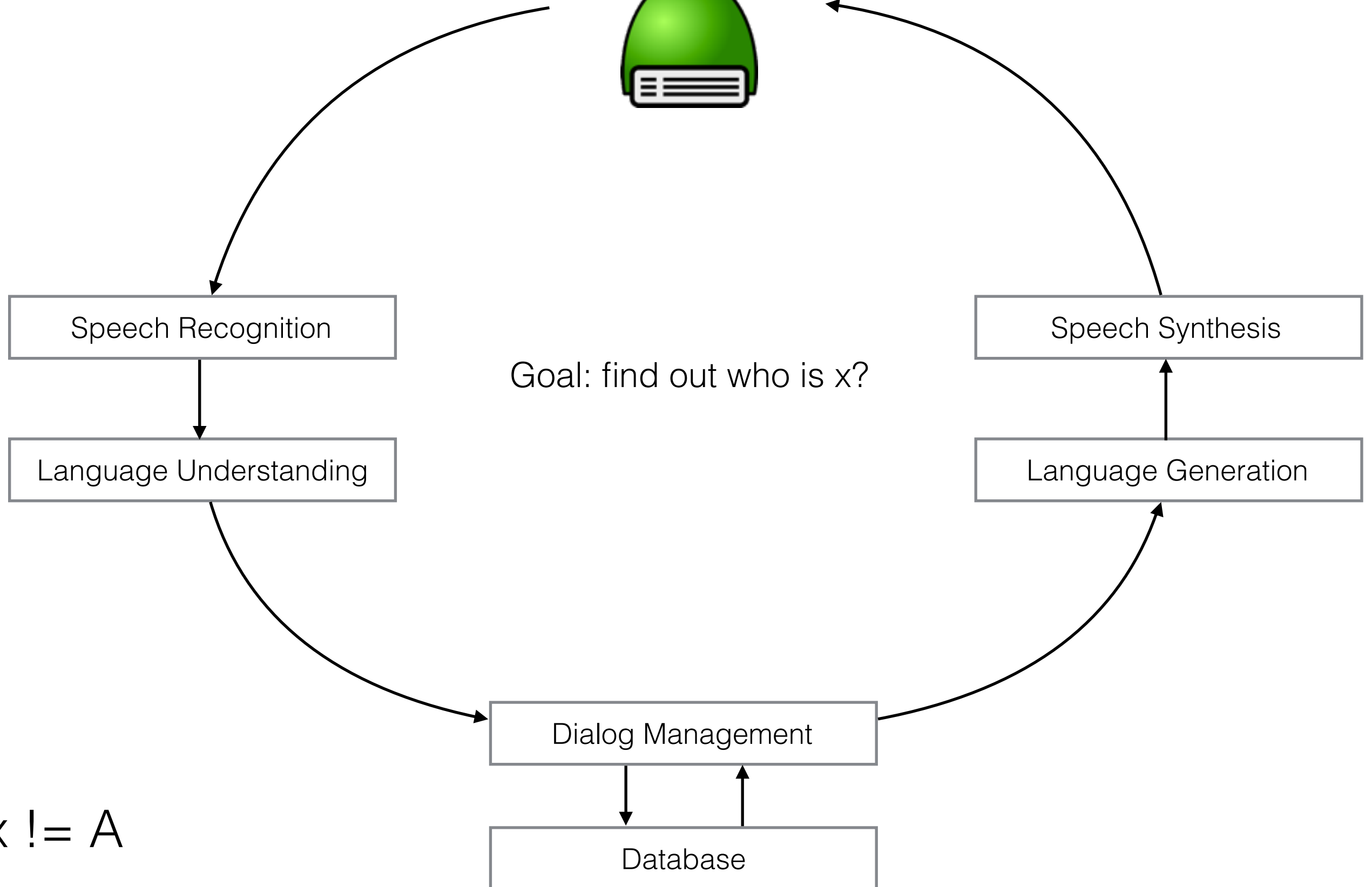
Speech Synthesis

Language Generation

Dialog Management

Database

$x \neq A$





Goal: find out who is x?

Speech Recognition

Language Understanding

Speech Synthesis

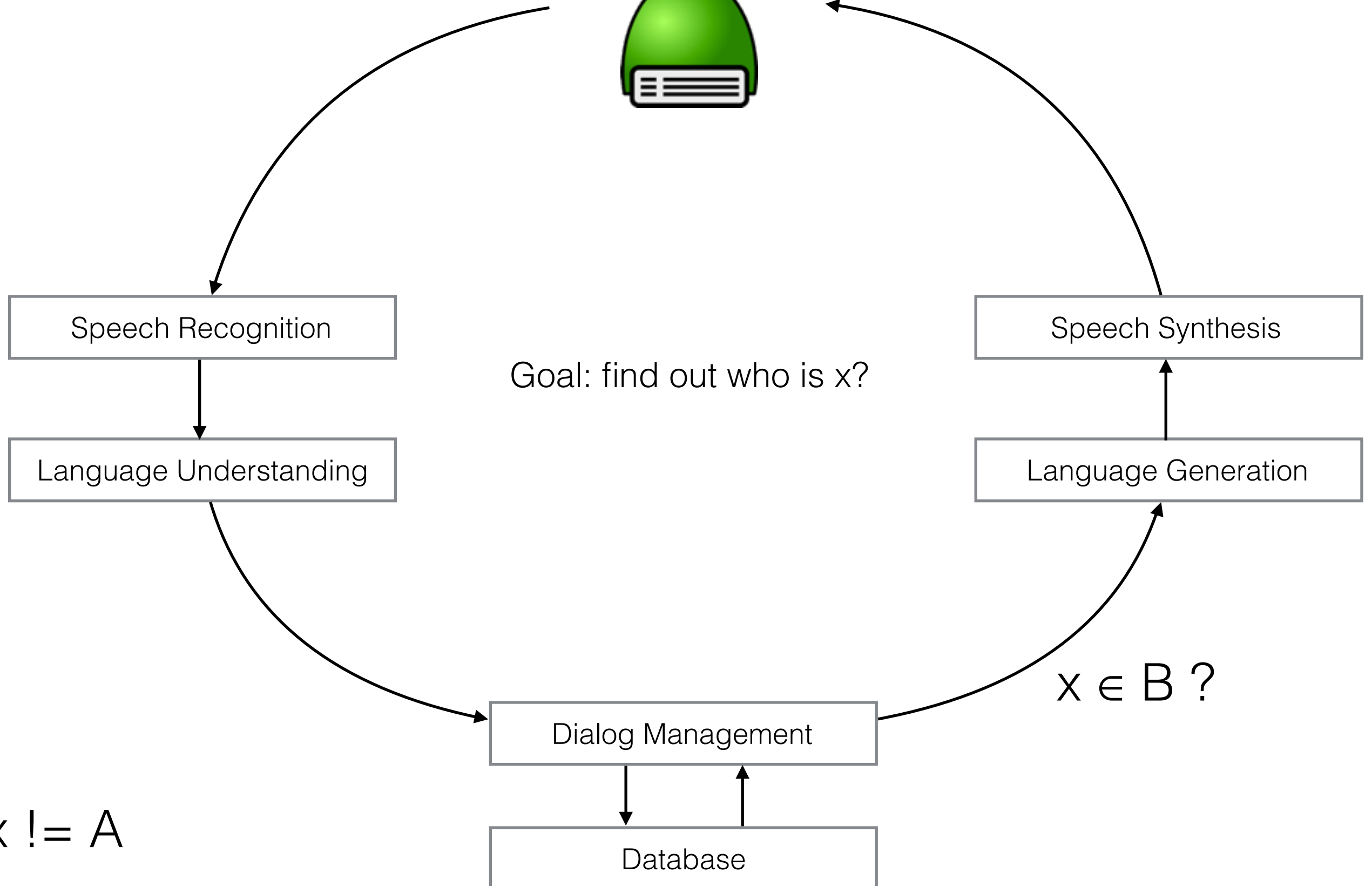
Language Generation

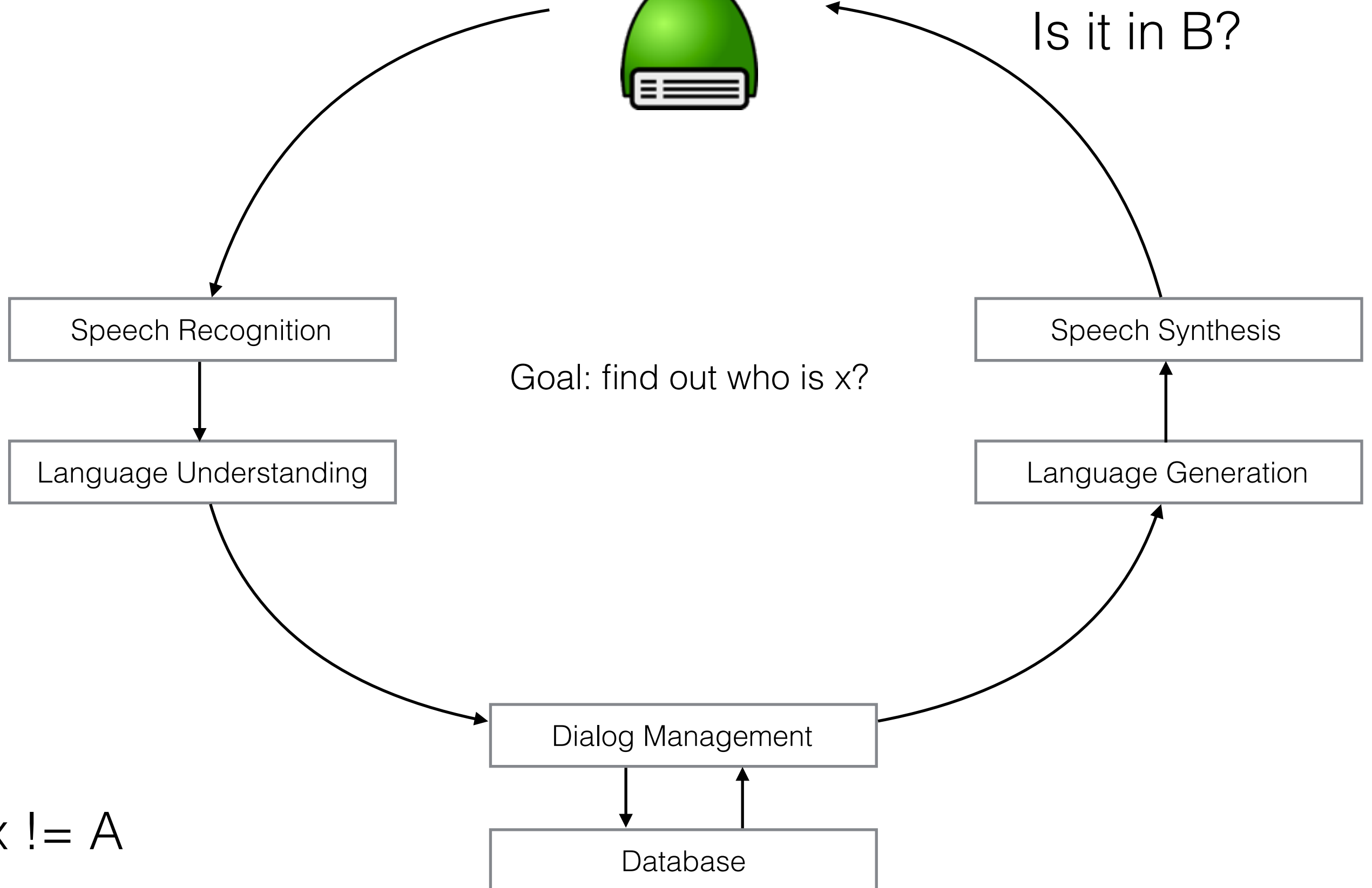
Dialog Management

Database

$x \neq A$

$x \in B ?$







No I don't think so.

Goal: find out who is x?

$x \neq A$

Speech Recognition

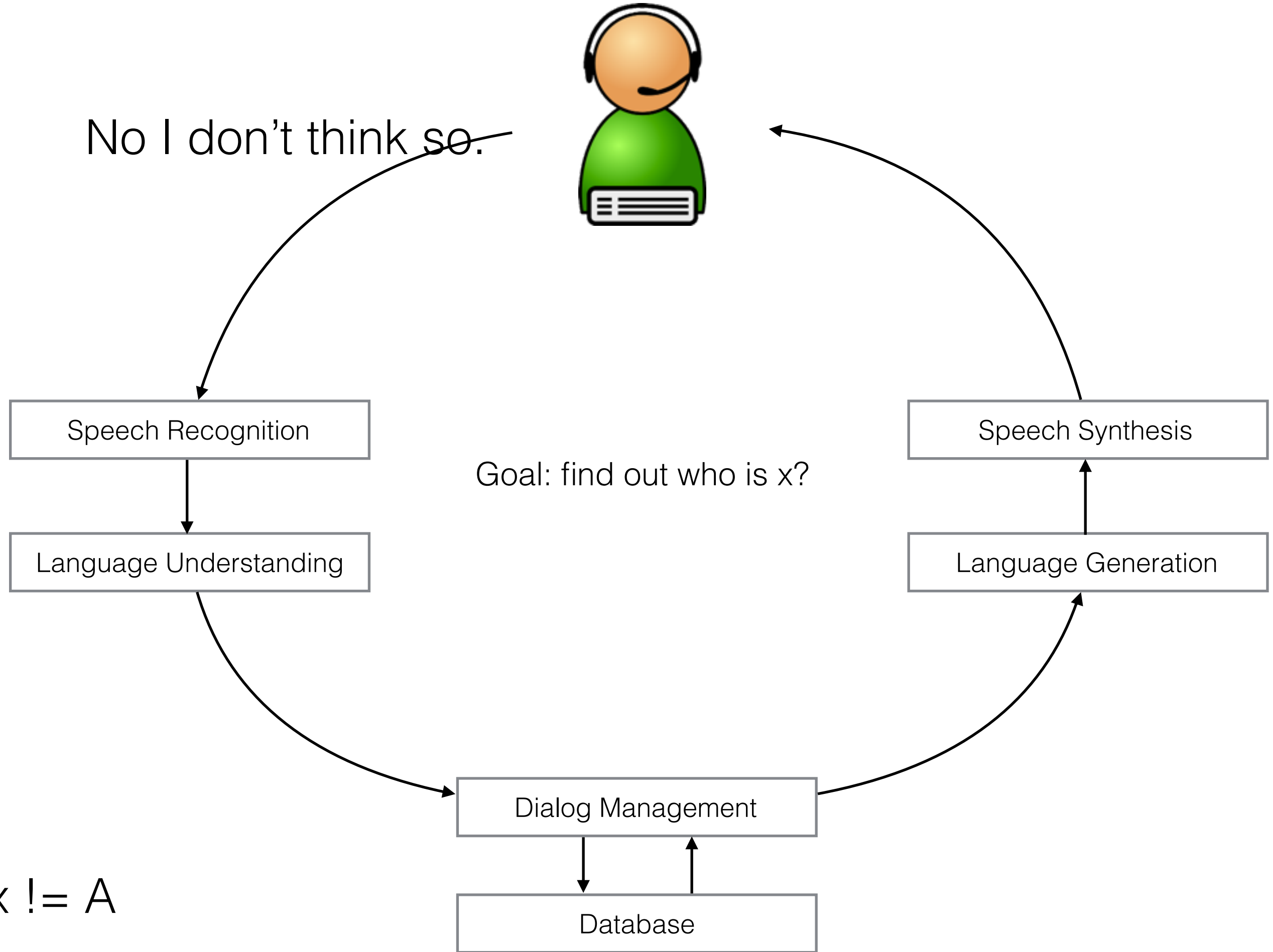
Language Understanding

Dialog Management

Database

Speech Synthesis

Language Generation





Goal: find out who is x?

Speech Recognition

Language Understanding

Speech Synthesis

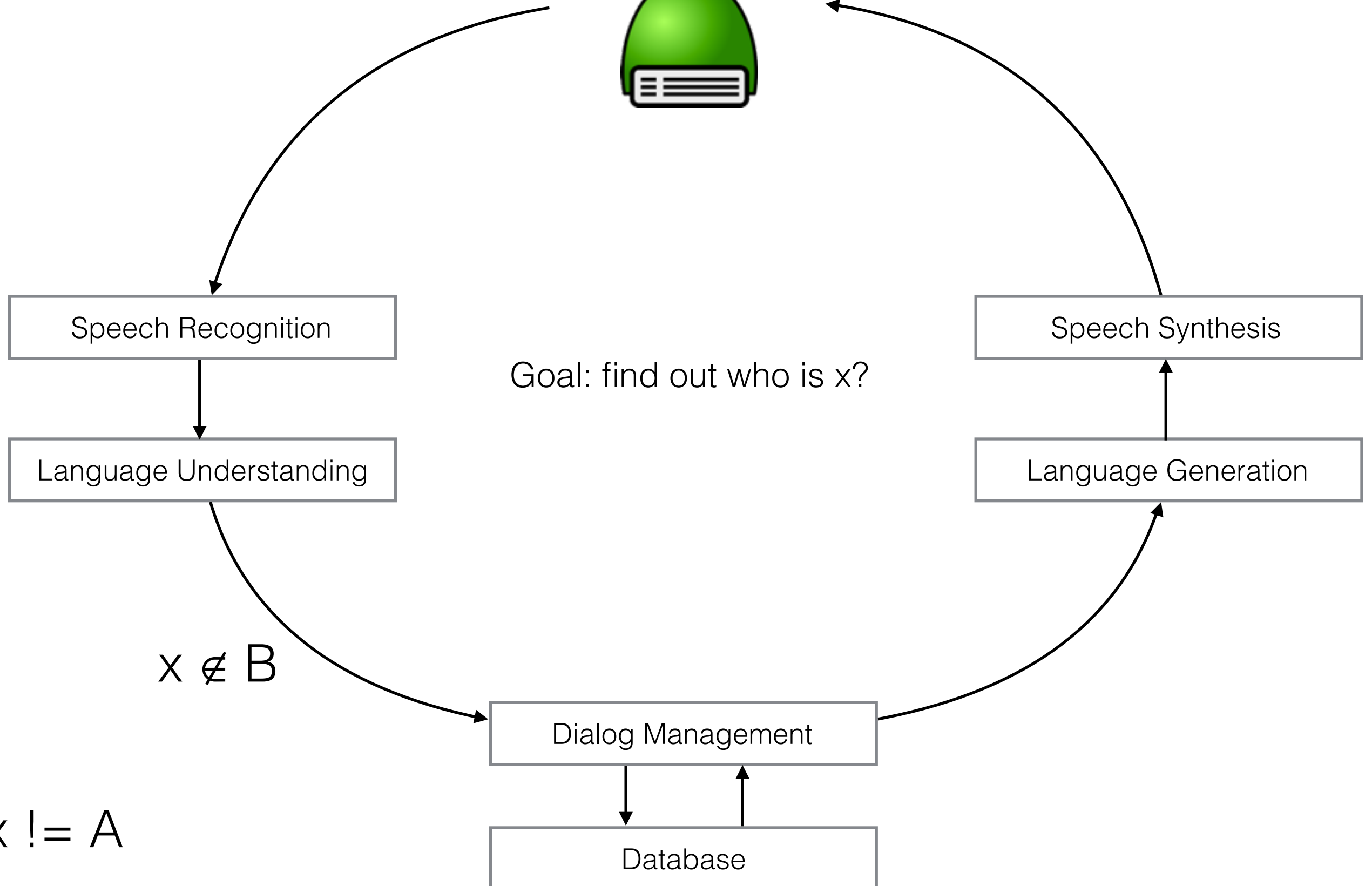
Language Generation

Dialog Management

Database

$x \notin B$

$x \neq A$





Goal: find out who is x?

Speech Recognition

Language Understanding

Speech Synthesis

Language Generation

Dialog Management

Database

$x \neq A$
 $x \notin B$

How it works?

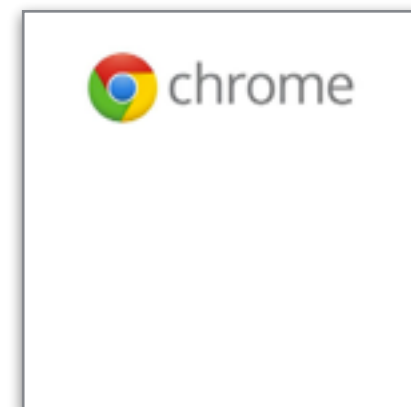
Technological perspective



How it works?

Technological perspective

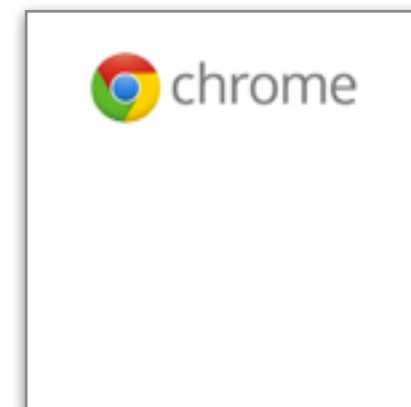
Browser



How it works?

Technological perspective

Browser



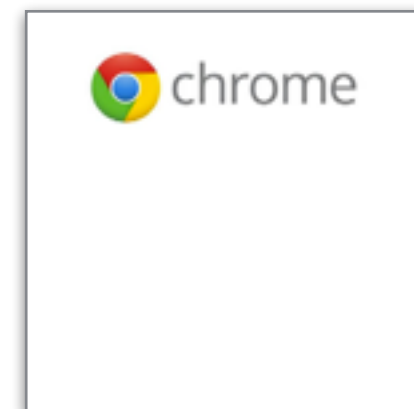
HTML/CSS

Javascript

How it works?

Technological perspective

Browser



HTML/CSS

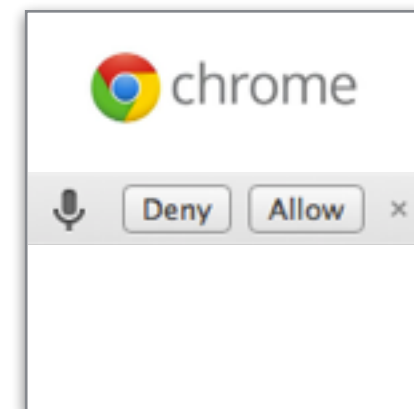
Javascript

Web Speech API

How it works?

Technological perspective

Browser



HTML/CSS

Javascript

Web Speech API

How it works?

Technological perspective

Browser



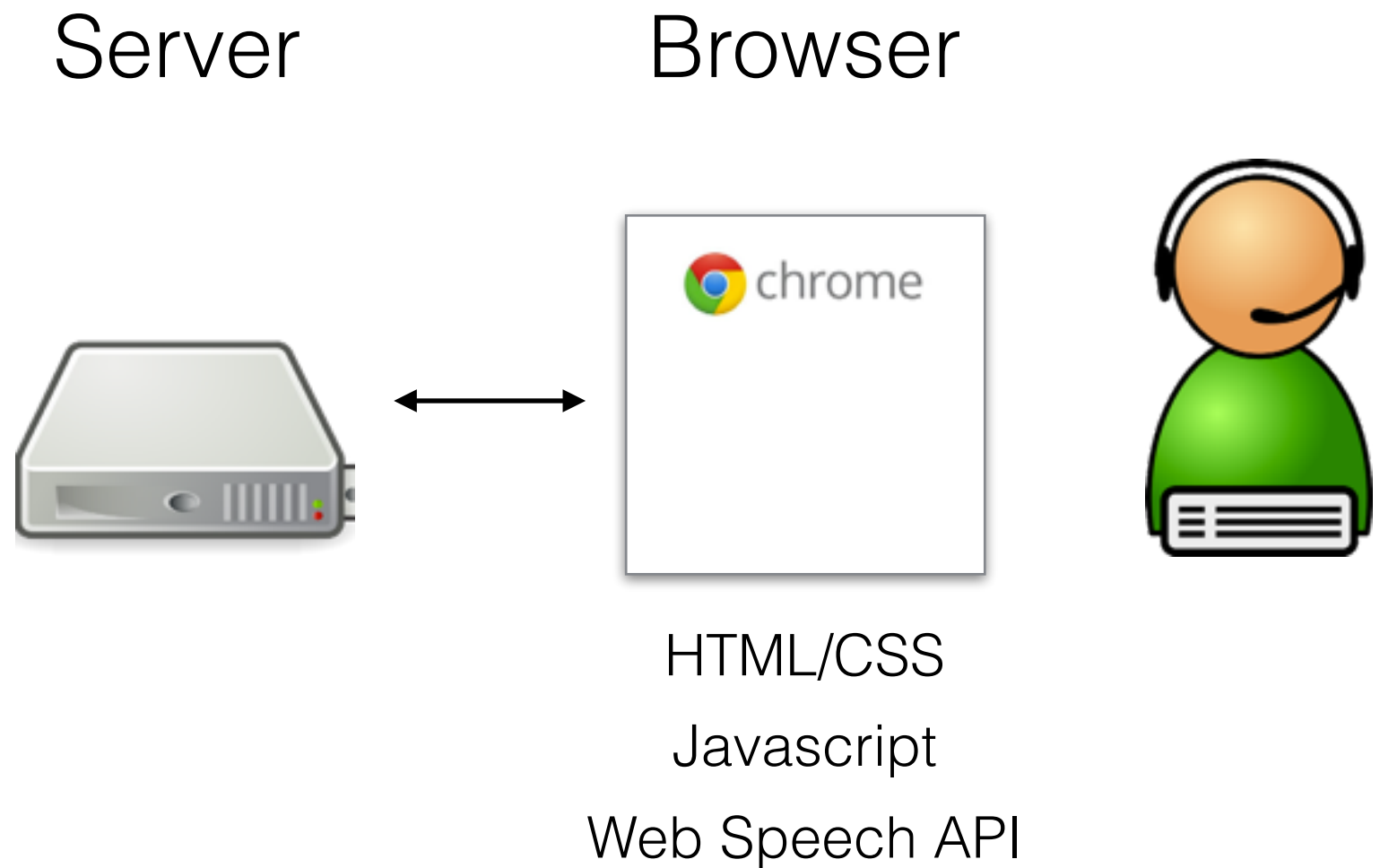
HTML/CSS

Javascript

Web Speech API

How it works?

Technological perspective

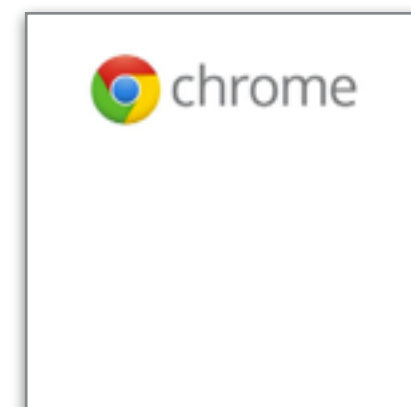


How it works?

Technological perspective

Server

Browser



NLTK, Language banks, ...

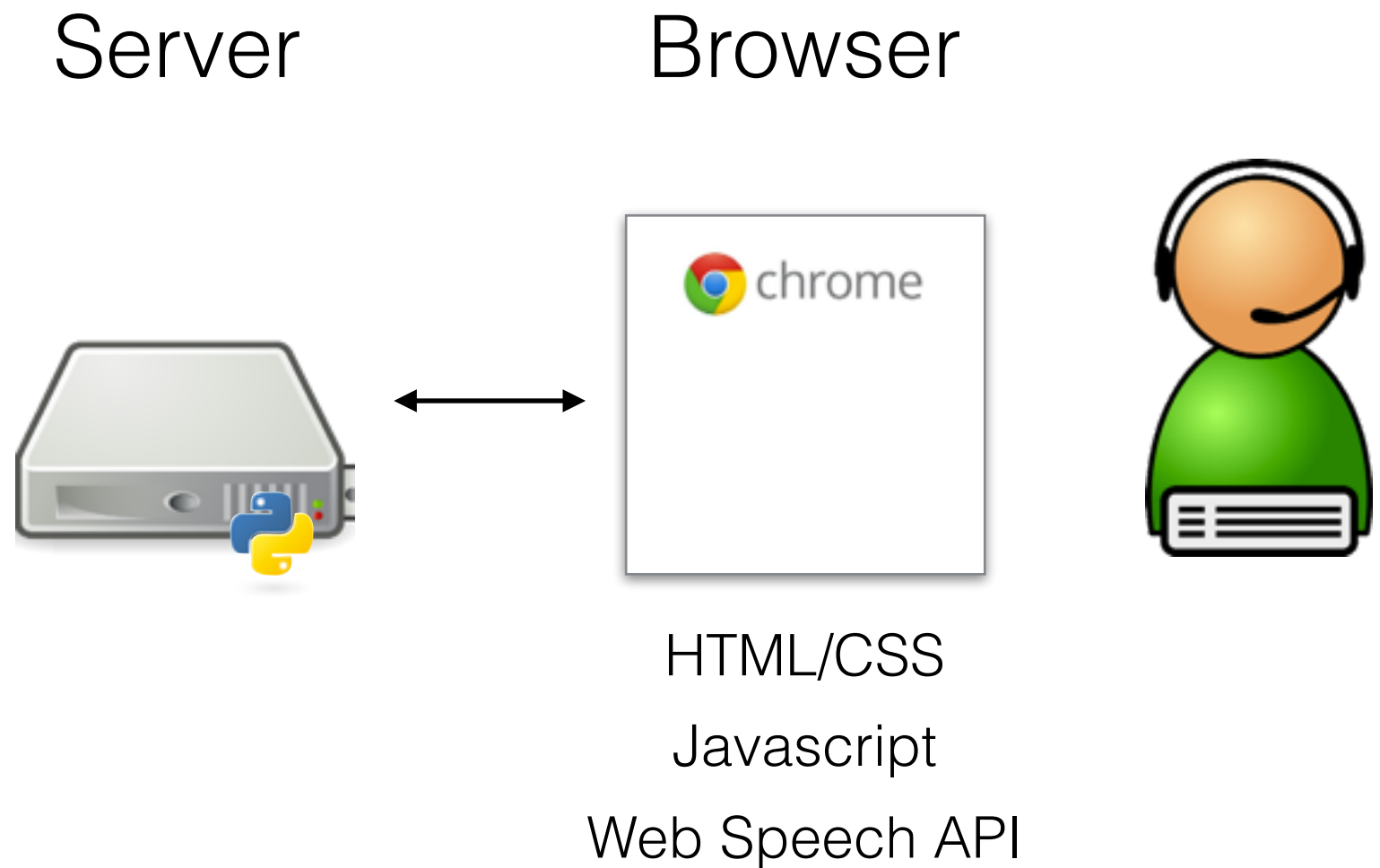
HTML/CSS

Javascript

Web Speech API

How it works?

Technological perspective



How it works?

Technological perspective

Knowledge base



Server



Browser



HTML/CSS
Javascript
Web Speech API



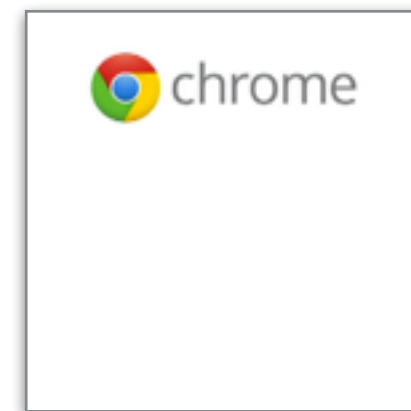
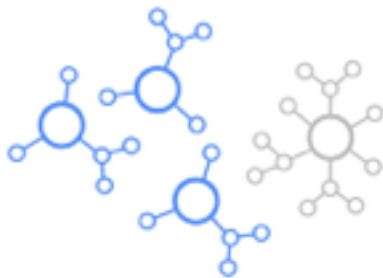
How it works?

Technological perspective

Knowledge base

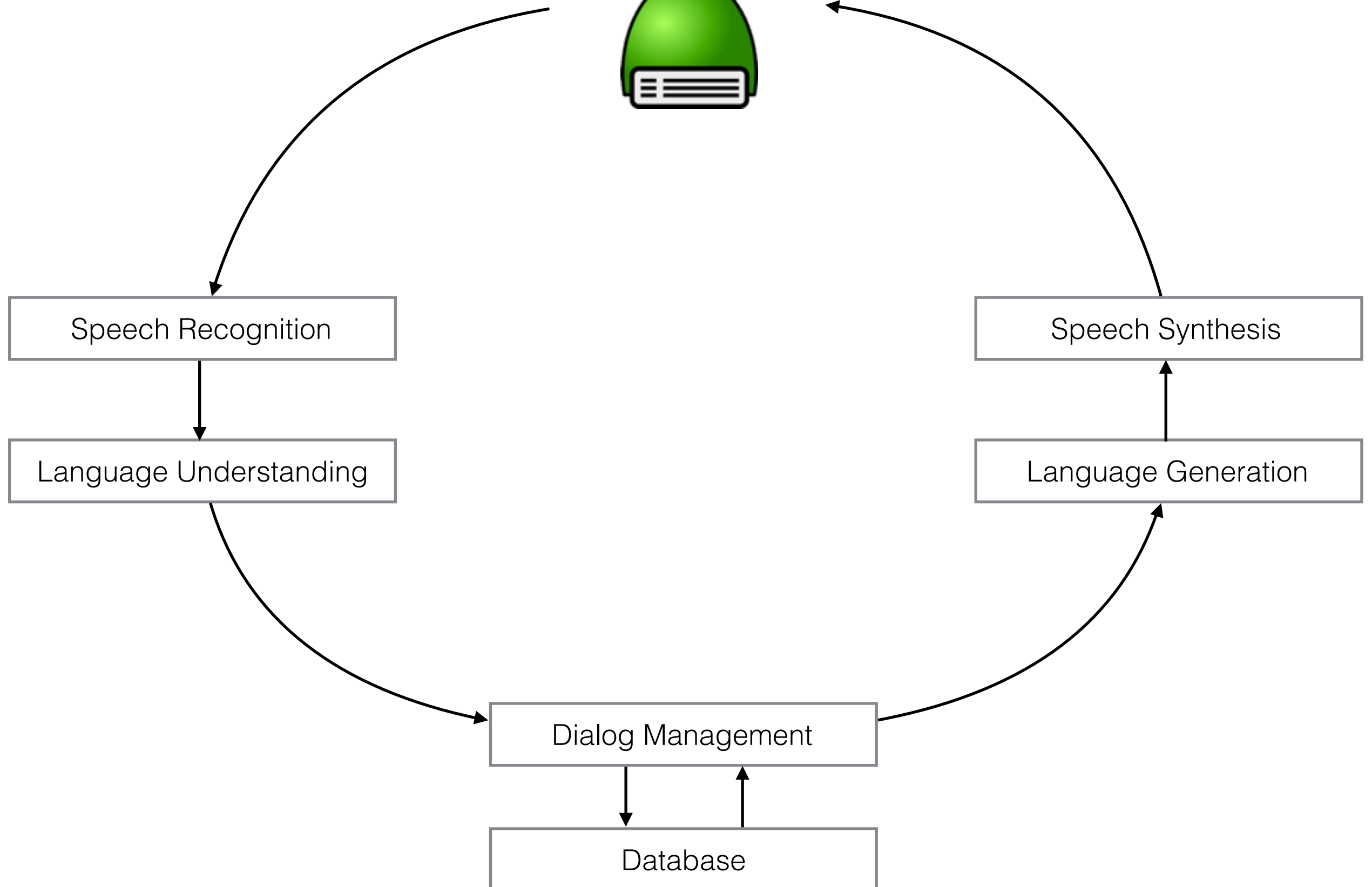
Server

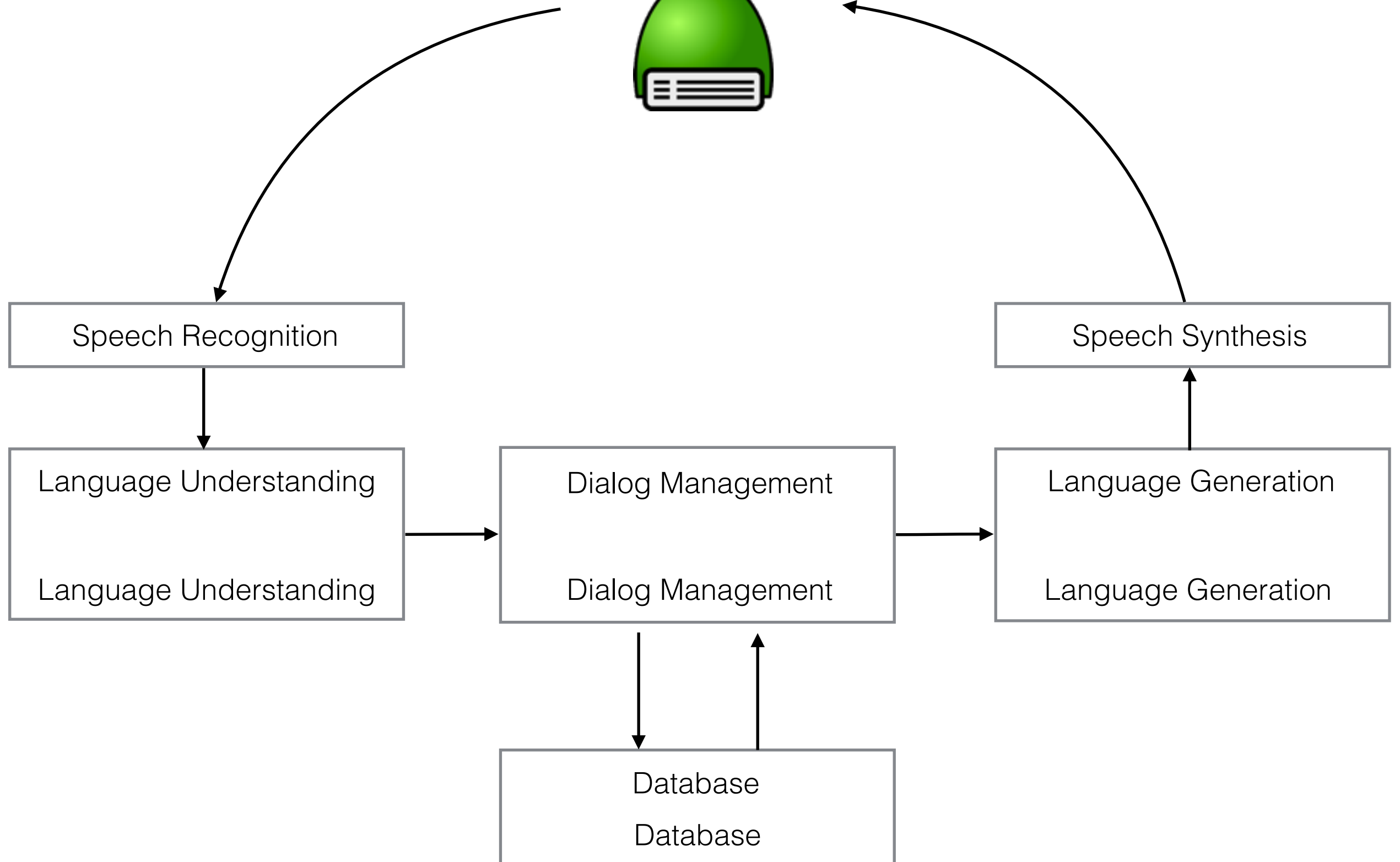
Browser



HTML/CSS
Javascript
Web Speech API









Client-side web application

Speech Recognition

Speech Synthesis

Language Understanding

Dialog Management

Language Generation

Language Understanding

Dialog Management

Language Generation

Database
Database





Client-side web application

Speech Recognition

Speech Synthesis

Language Understanding

Language Understanding

Dialog Management

Dialog Management

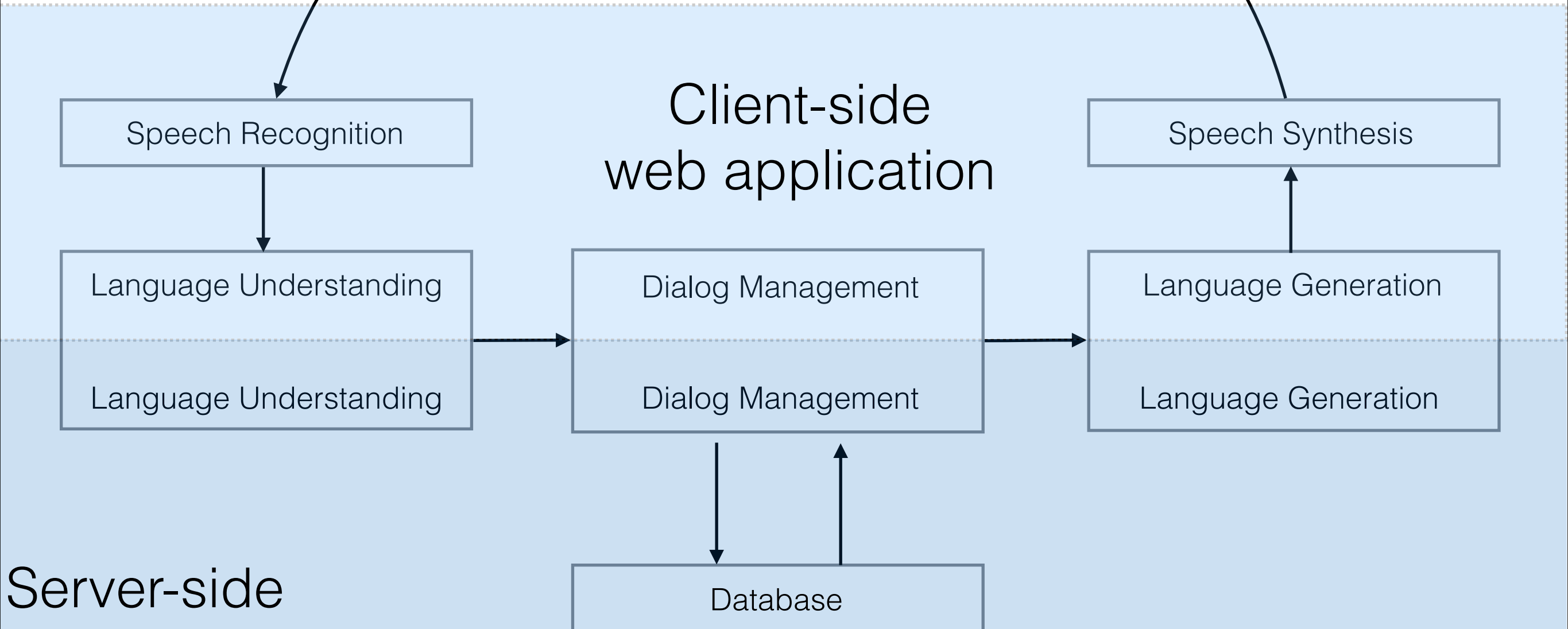
Language Generation

Language Generation

Database

Database

Server-side





Client-side web application

Speech Recognition

Speech Synthesis

Language Understanding

Language Understanding

Dialog Management

Dialog Management

Language Generation

Language Generation

Database

Database

Server-side

Knowledge base

How it works?

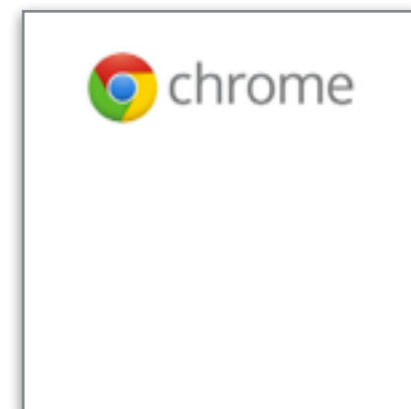
Process perspective



How it works?

Process perspective

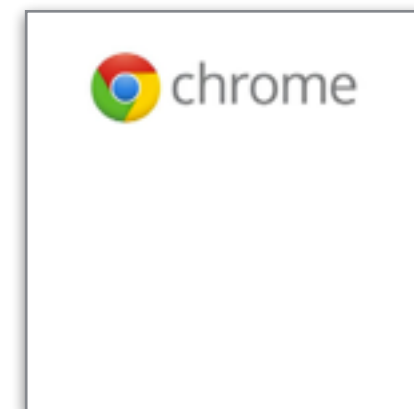
Browser



How it works?

Process perspective

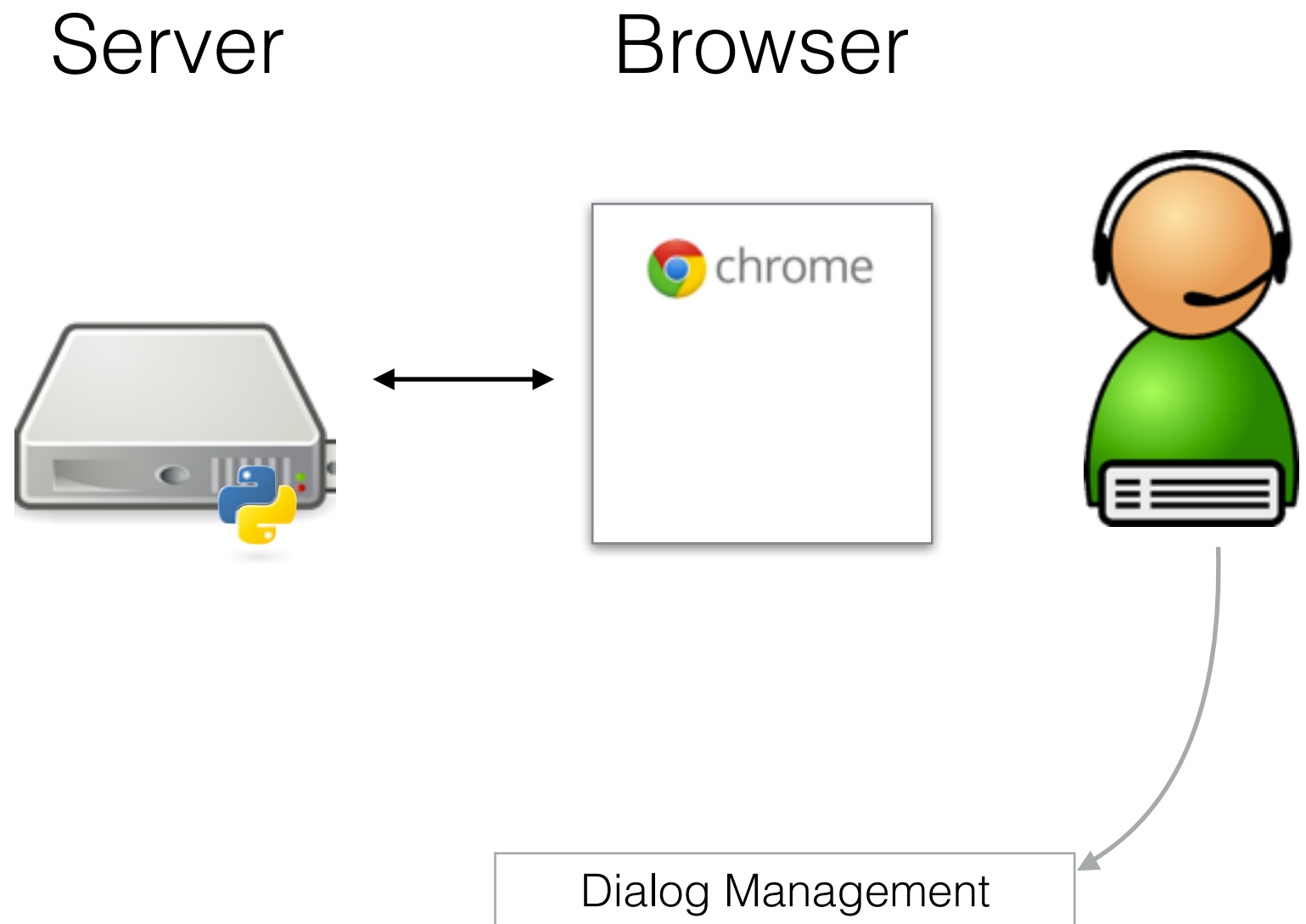
Browser



Dialog Management

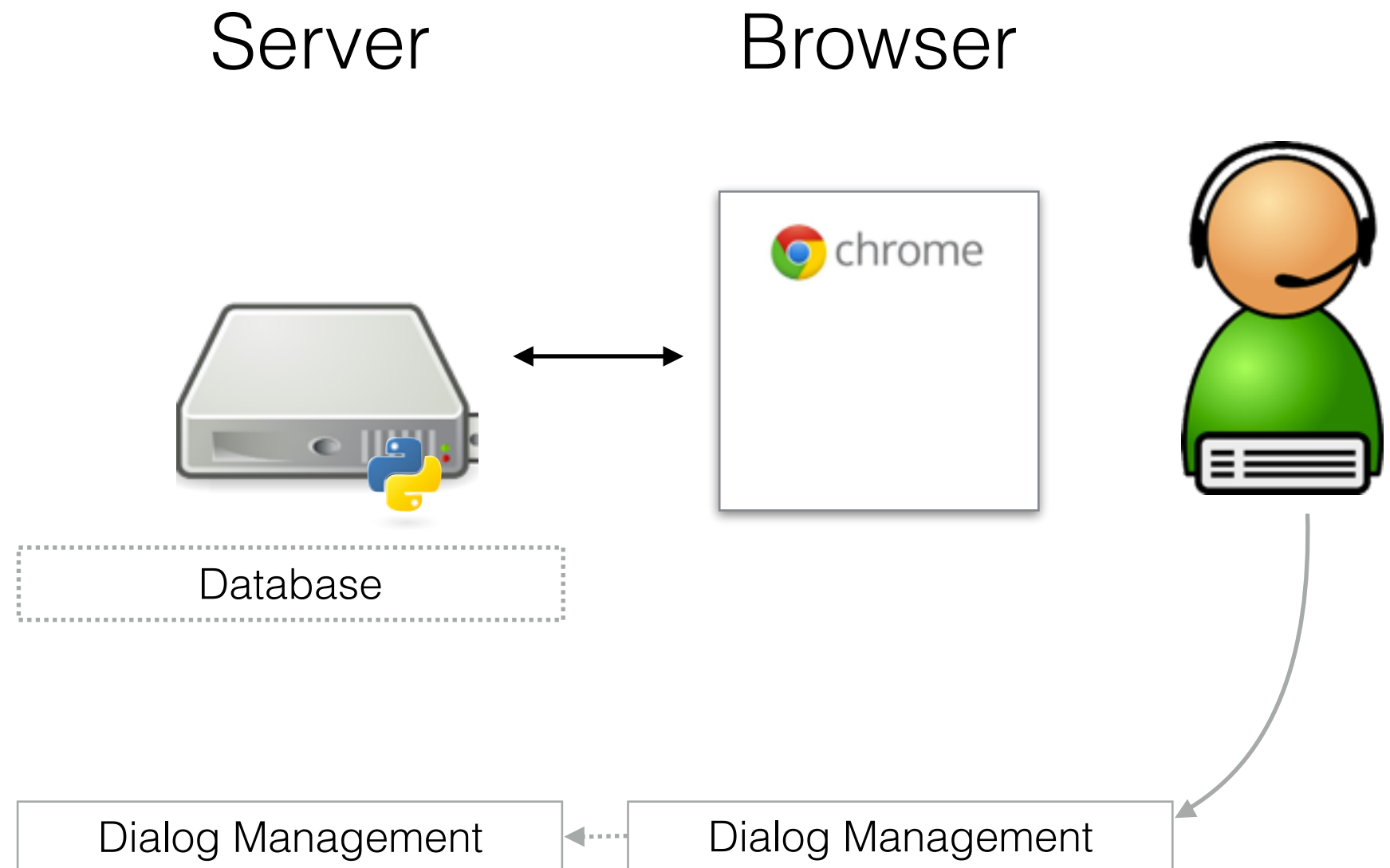
How it works?

Process perspective



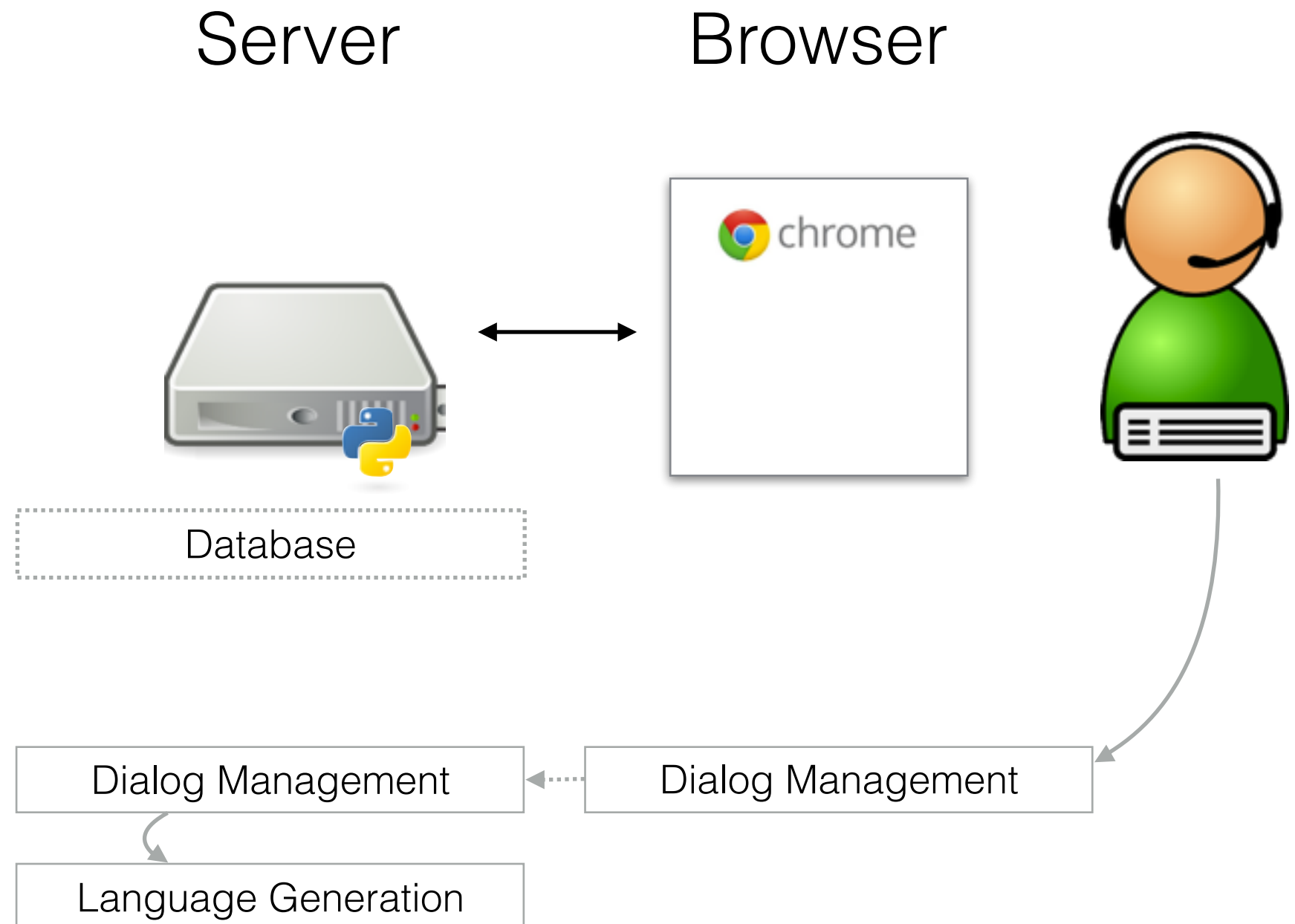
How it works?

Process perspective



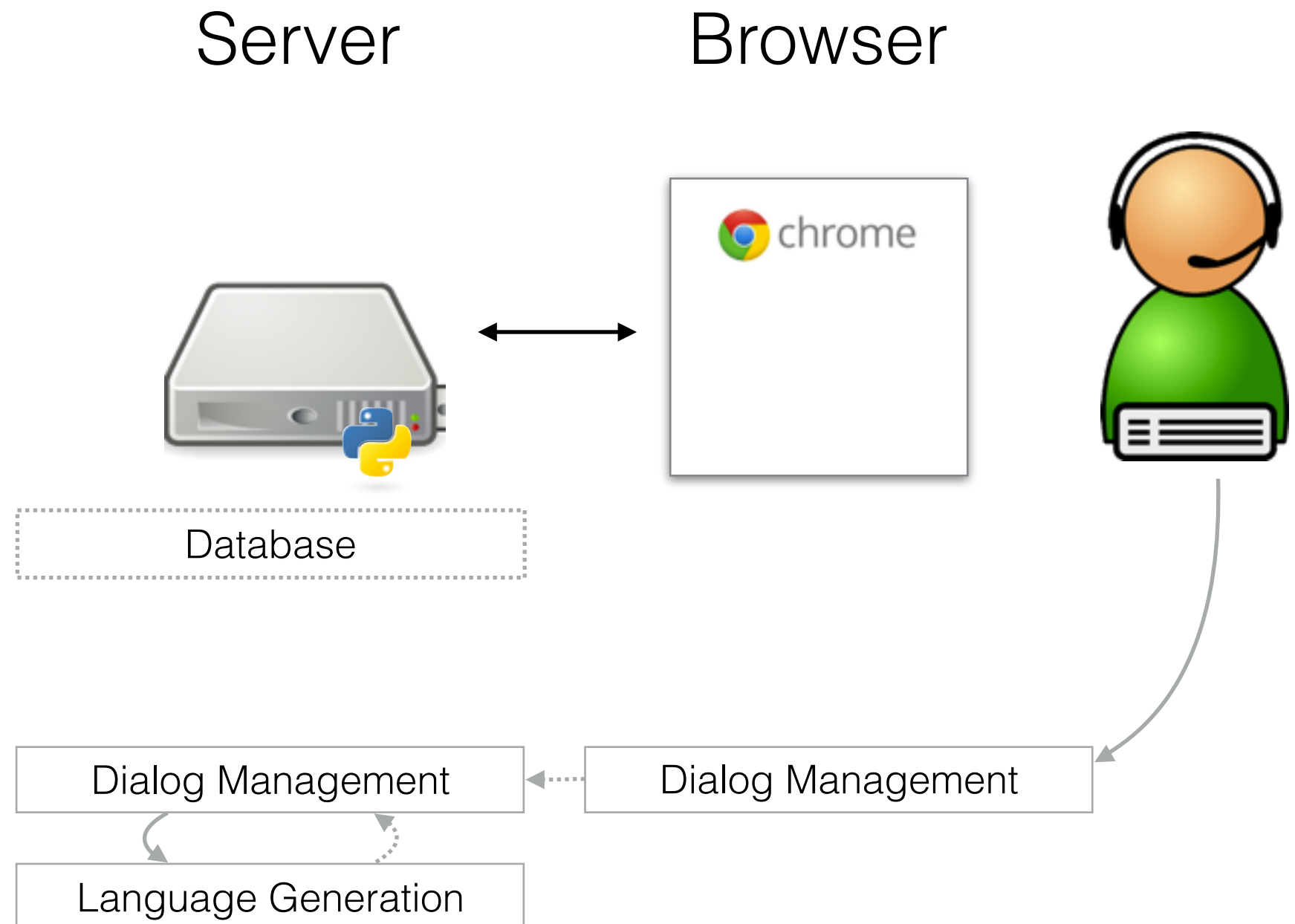
How it works?

Process perspective



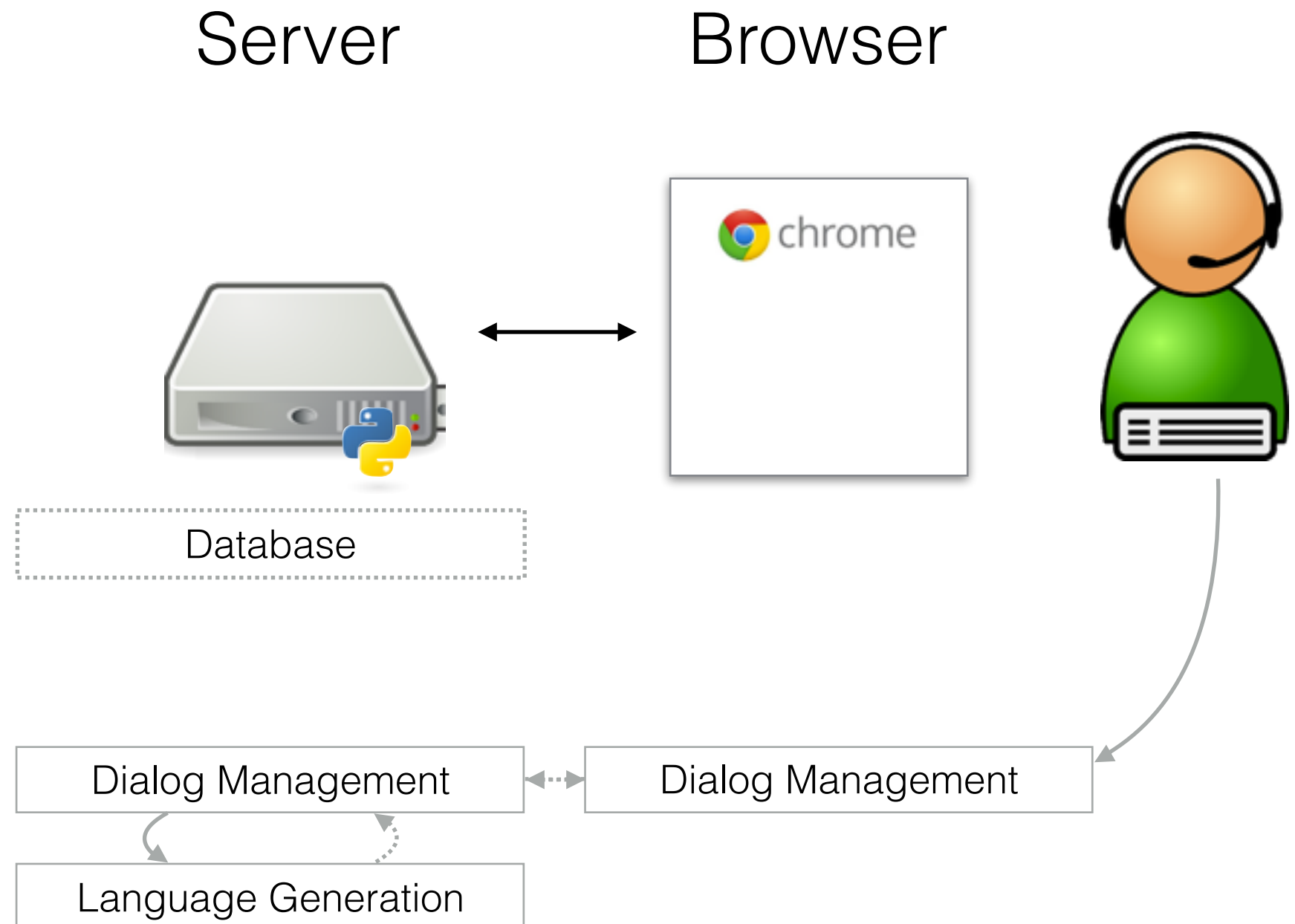
How it works?

Process perspective



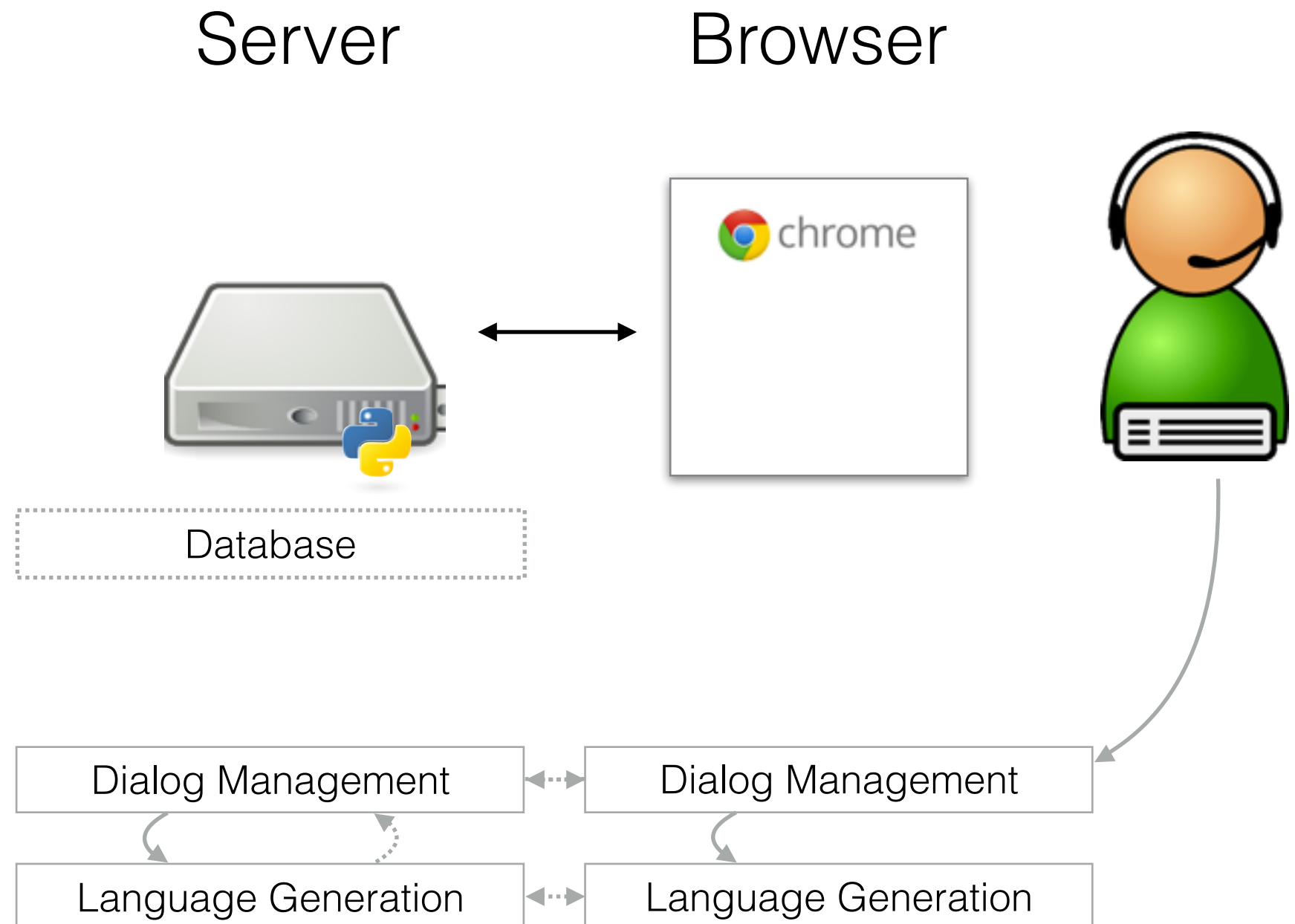
How it works?

Process perspective



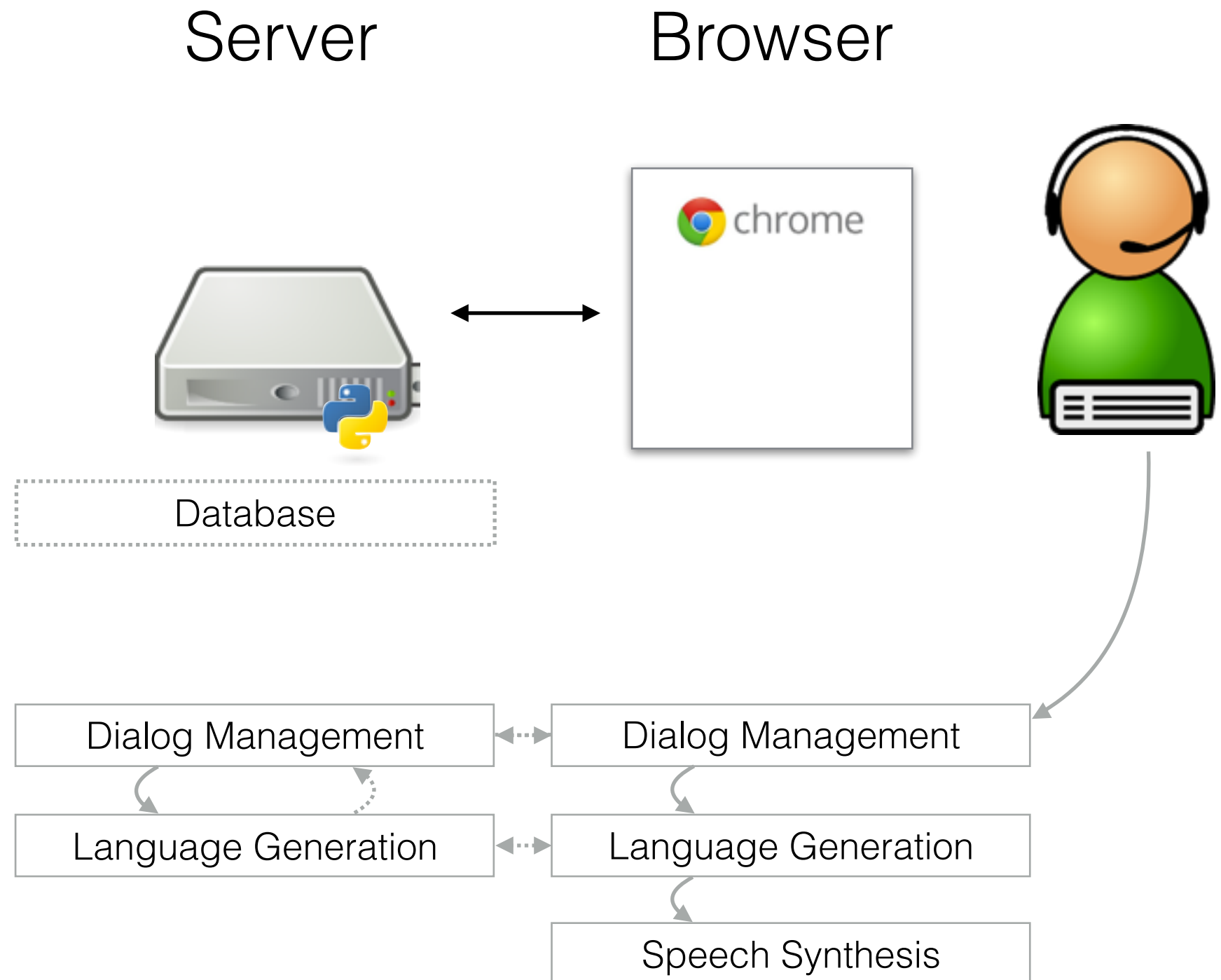
How it works?

Process perspective



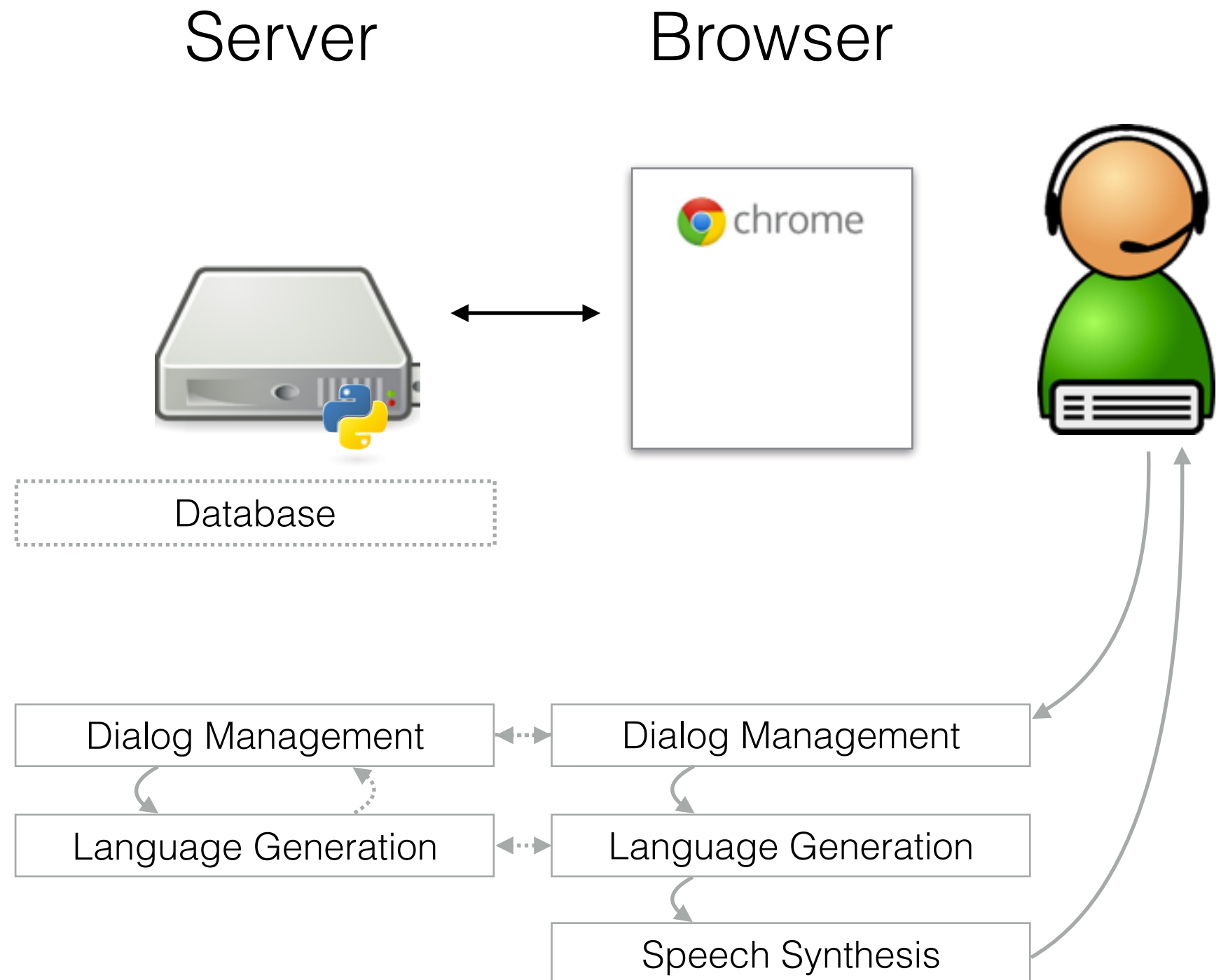
How it works?

Process perspective



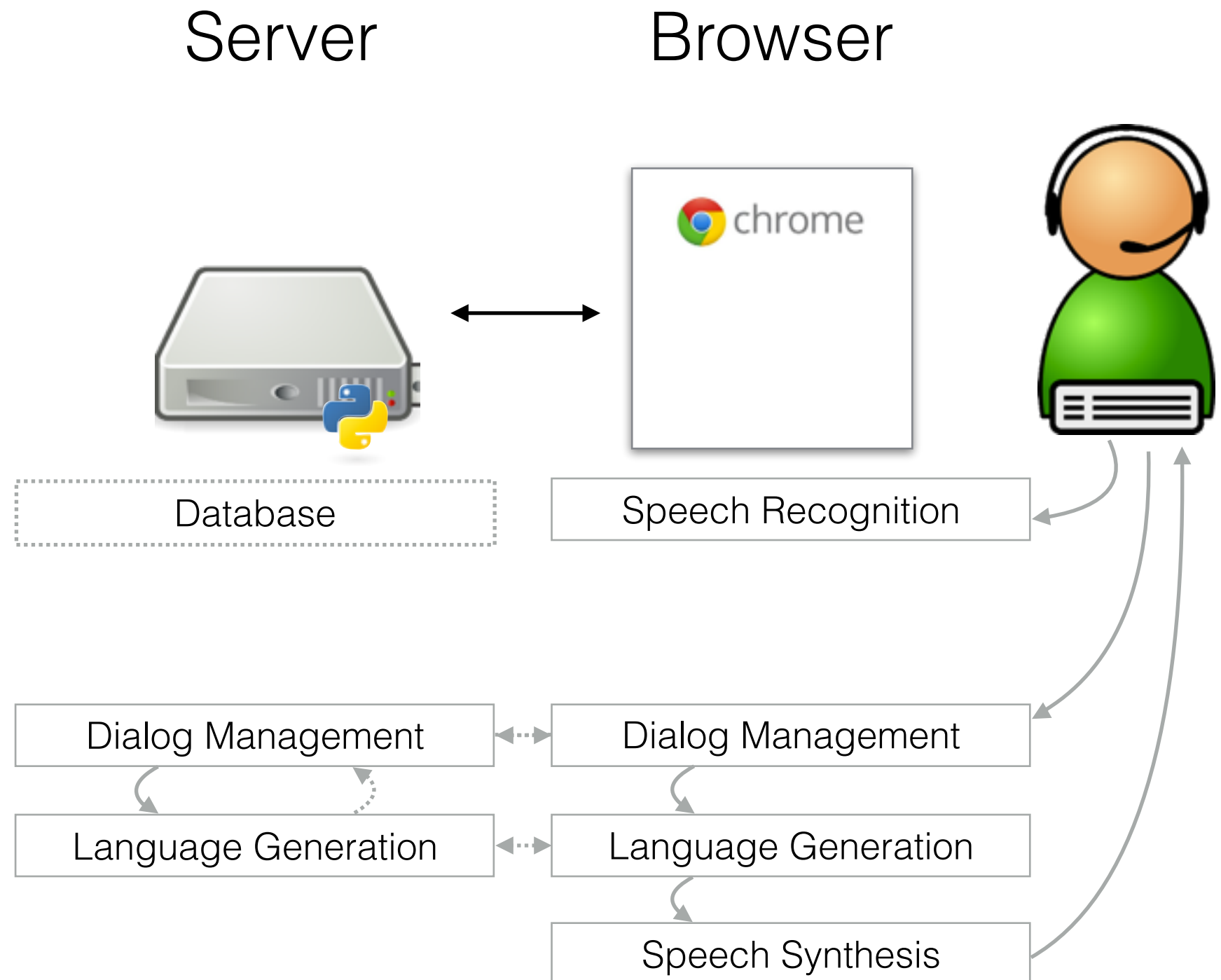
How it works?

Process perspective



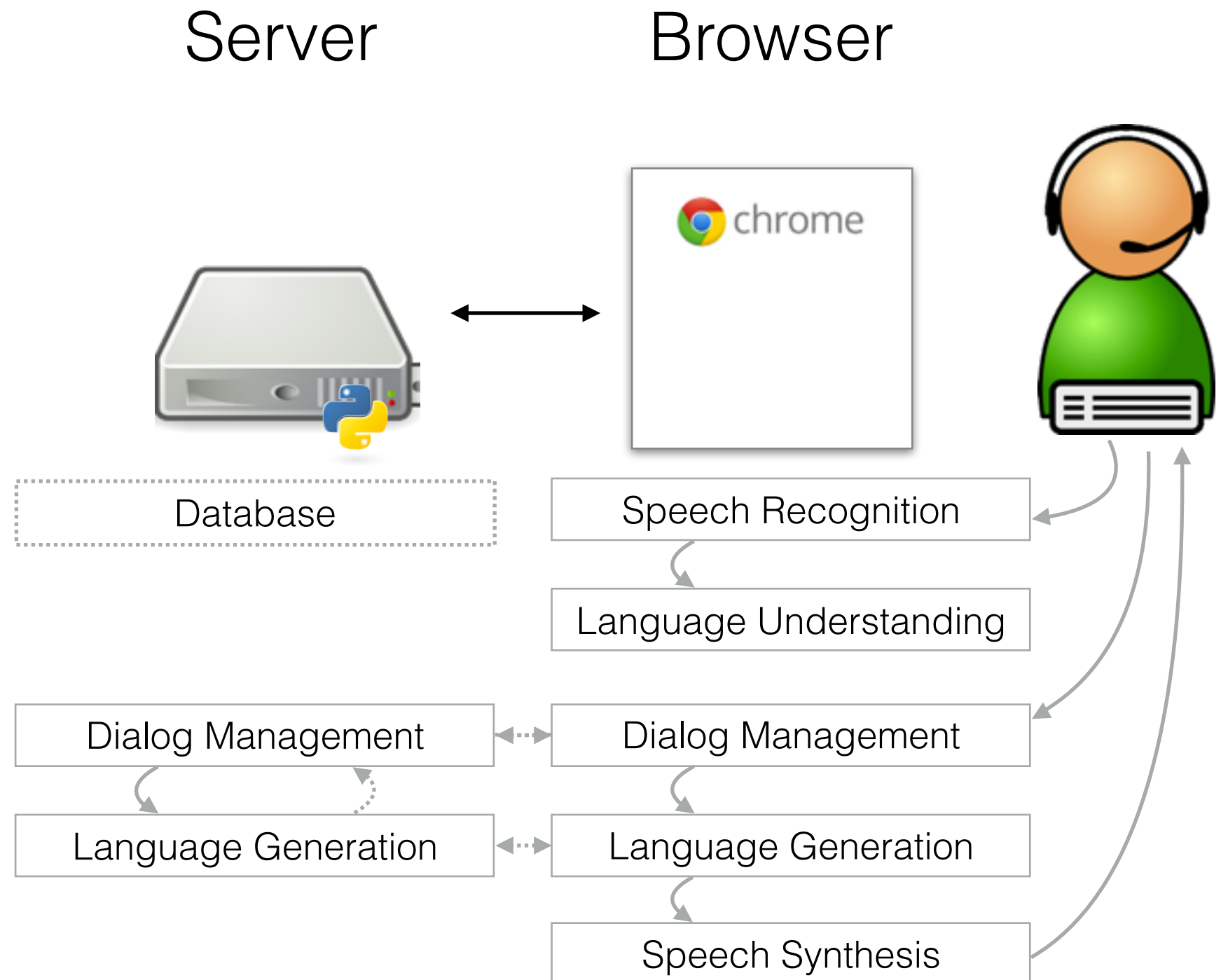
How it works?

Process perspective



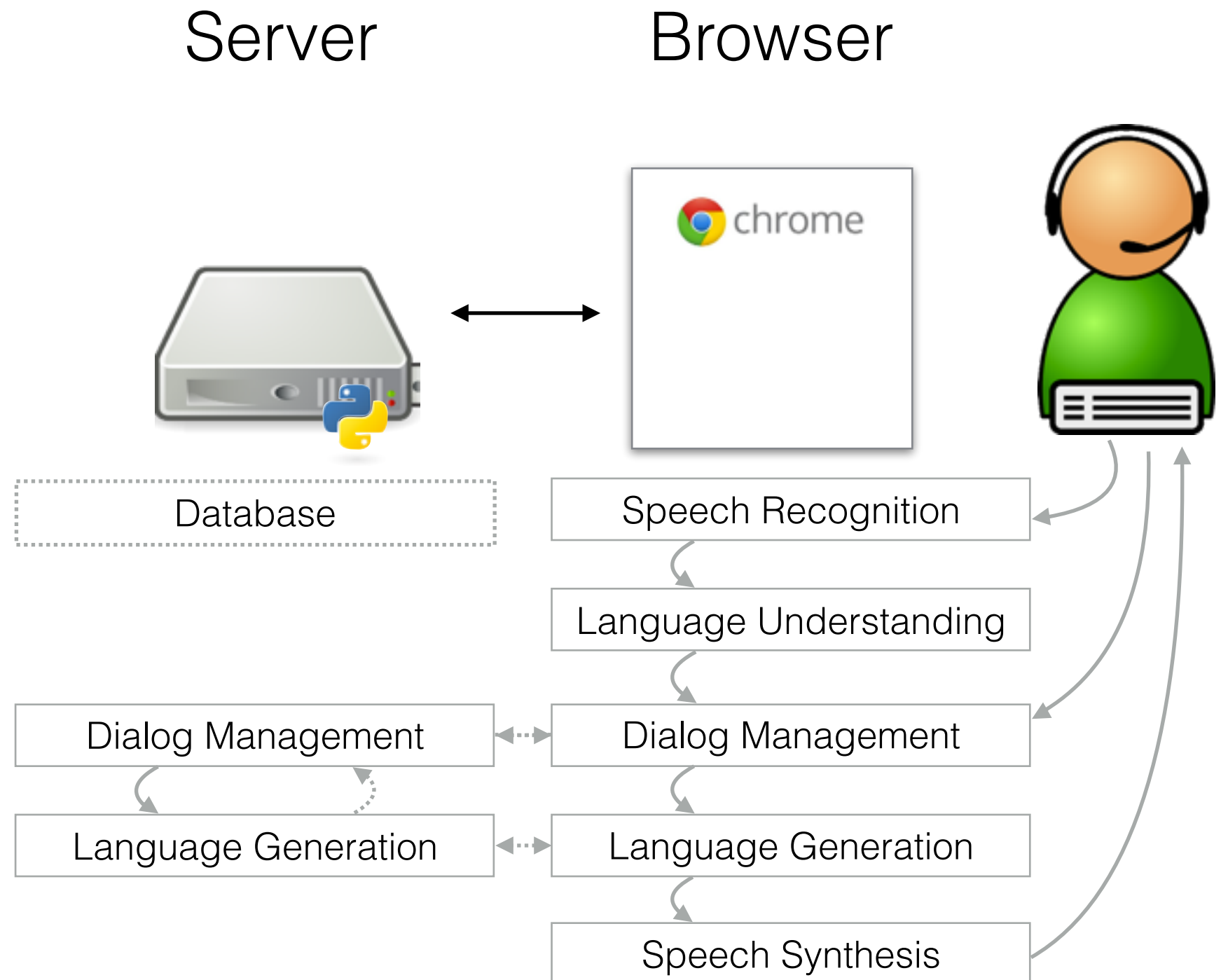
How it works?

Process perspective



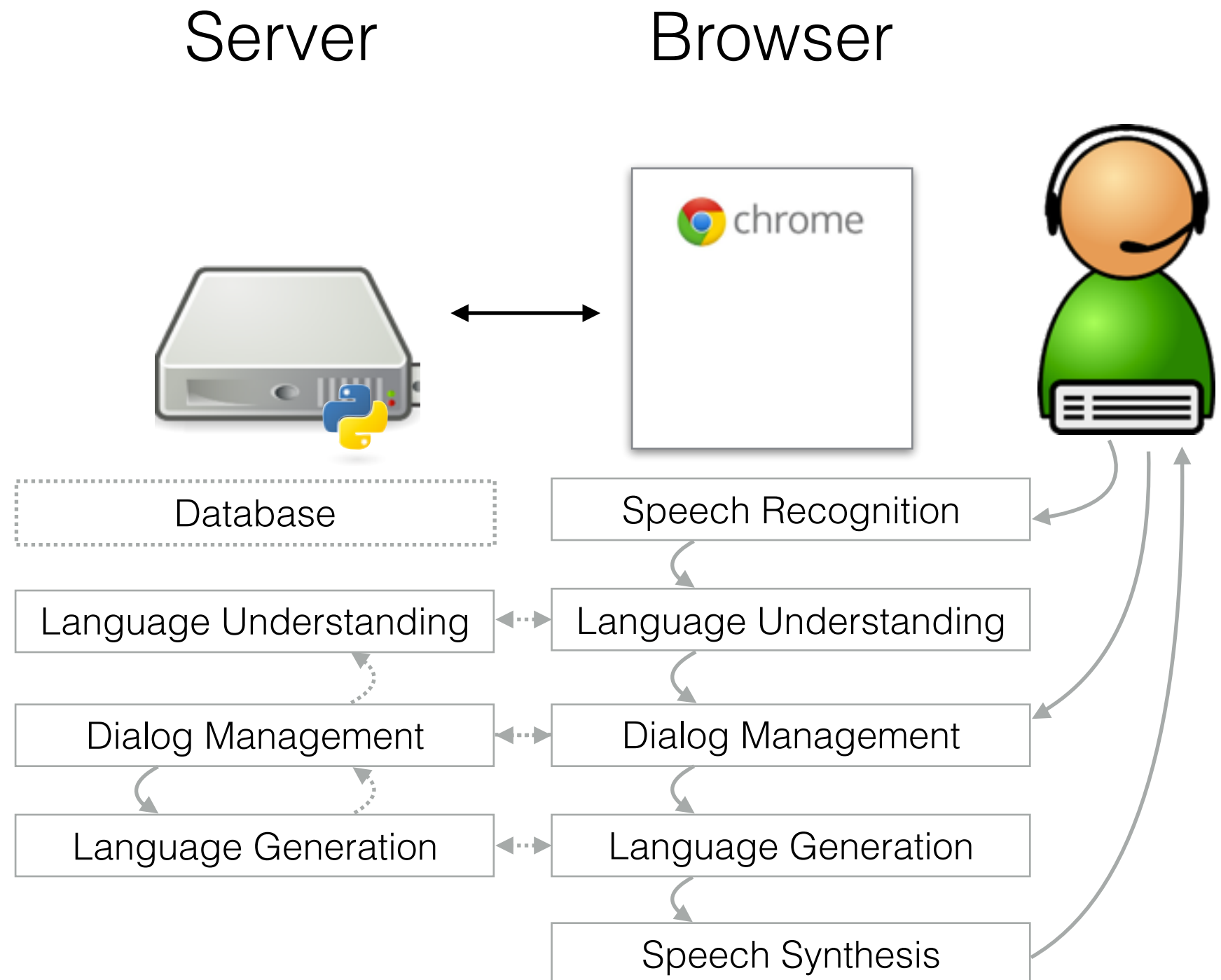
How it works?

Process perspective



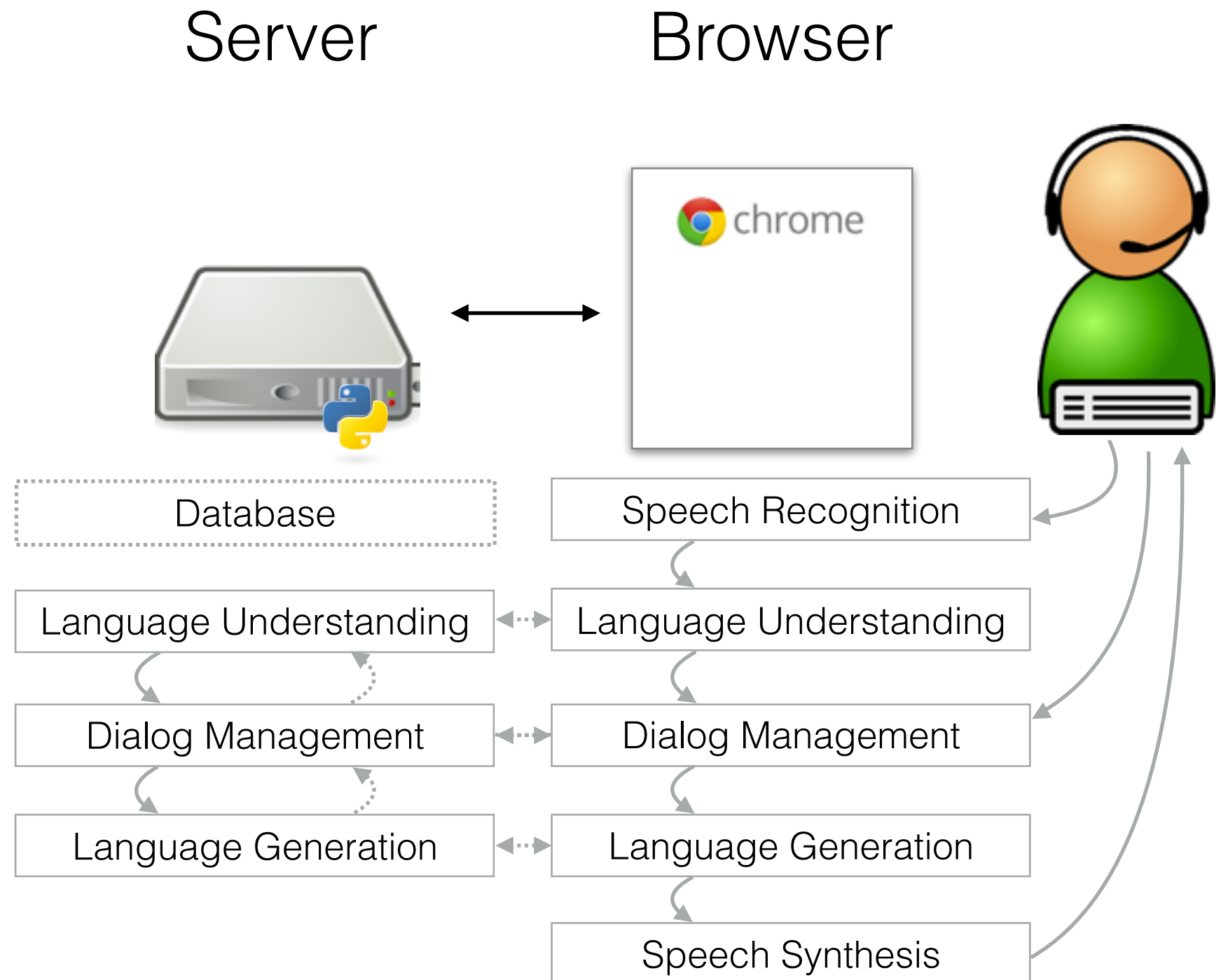
How it works?

Process perspective



How it works?

Process perspective



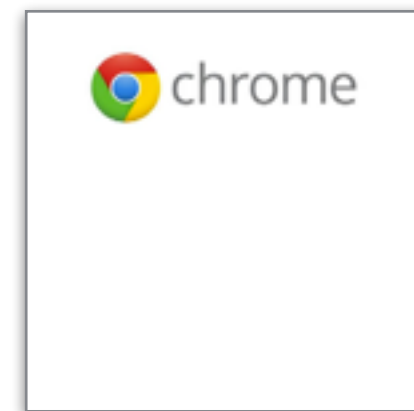
How it works?

Process perspective

Knowledge base

Server

Browser



Database

Language Understanding

Dialog Management

Language Generation

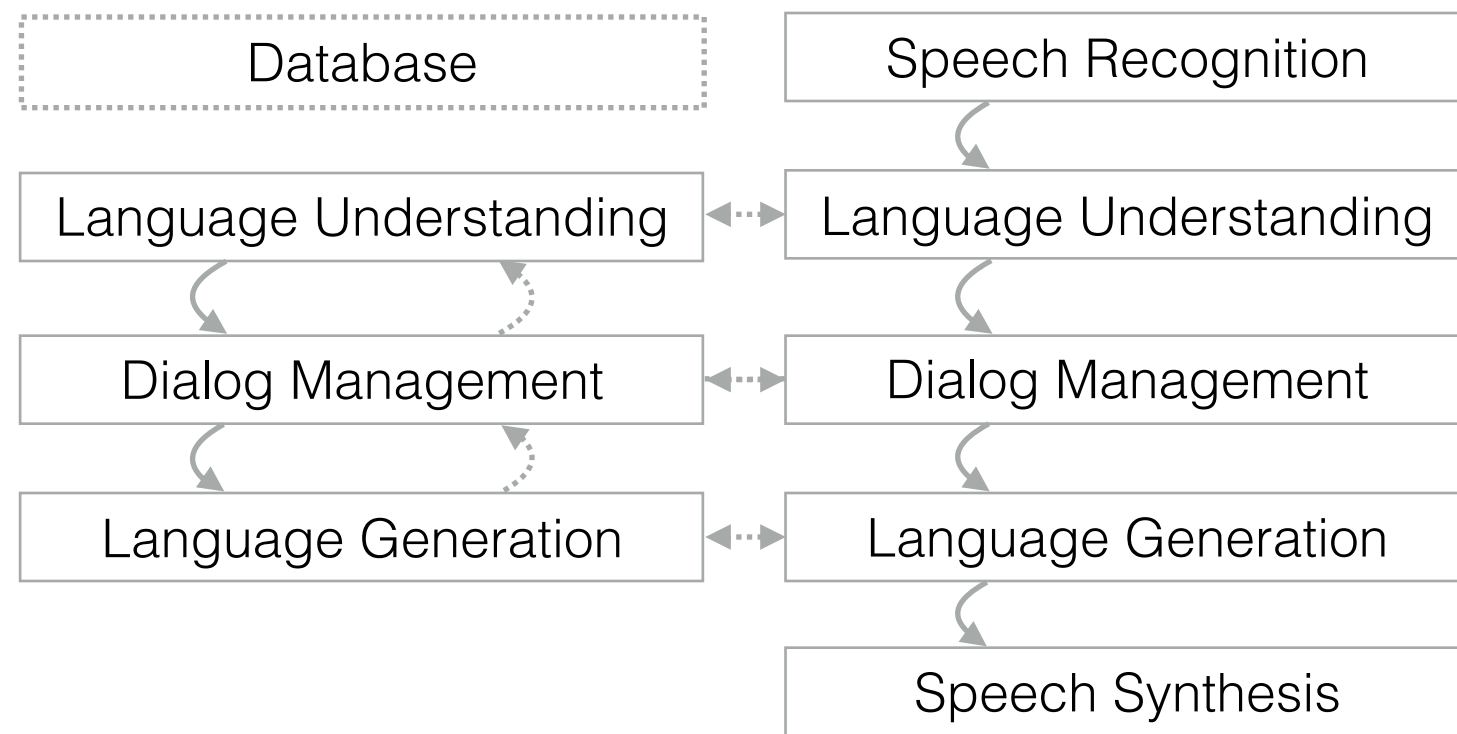
Speech Recognition

Language Understanding

Dialog Management

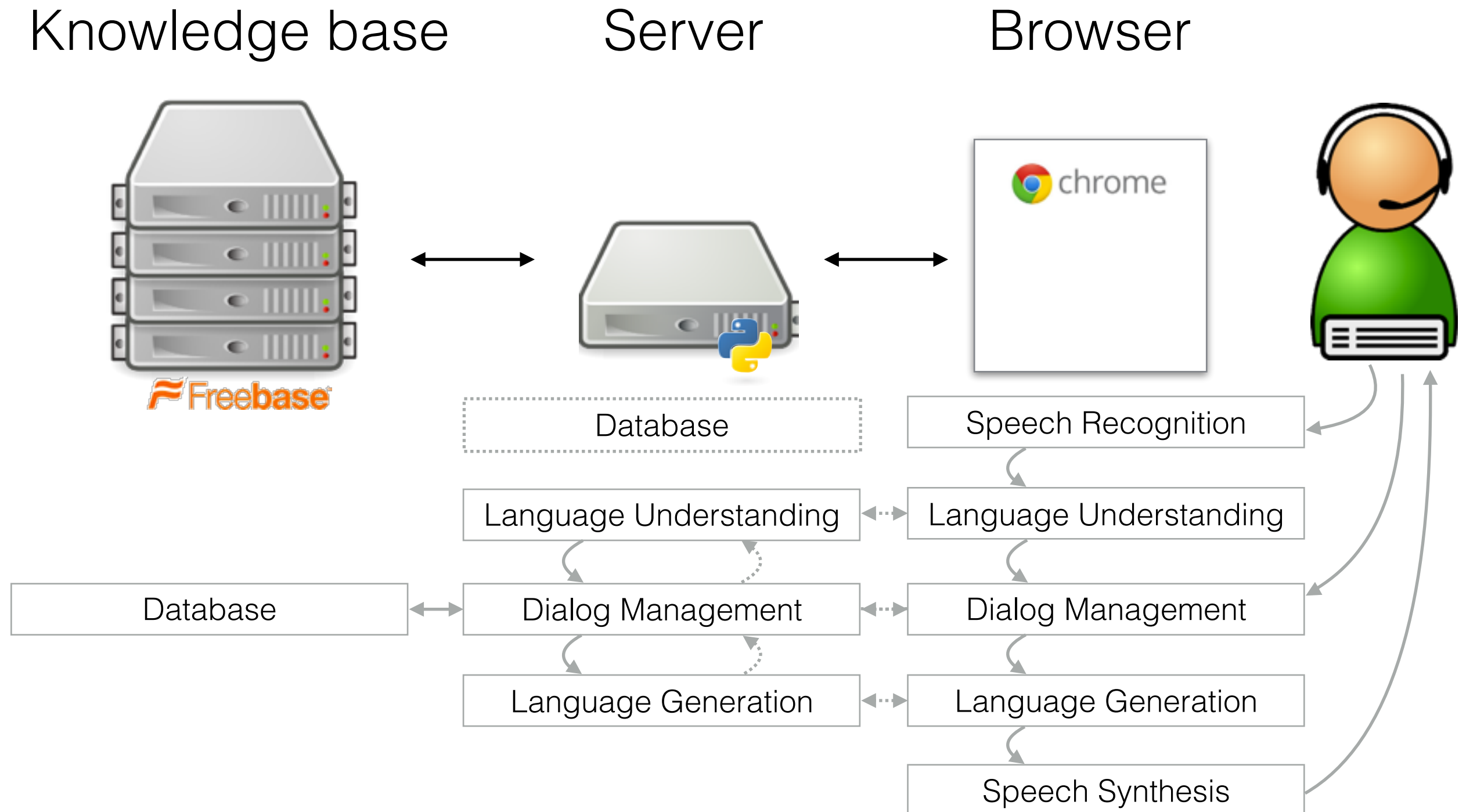
Language Generation

Speech Synthesis



How it works?

Process perspective



Technical perspective

- Browser (Javascript: Web speech API, jQuery)
- Server (Python: flask, urllib)
- Knowledge base (FreeBase API)

Wizard of the Oz

[Explanation]

dirty programming

Javascript introduction

- Javascript as a scripting language for HTML.
- You should know HTML and CSS to working with this javascript.
- Browser runs the code while downloading the `<script>` tag.
- You can load external files `<script src="myscript.js"></script>`

```
<!DOCTYPE html>
<html>
  <body>
    <h1>Javascript intro test</h1>
    <script>
      document.write("<p>write something on screen</p>");
    </script>
  </body>
</html>
```

Javascript introduction

What is DOM?

When a web page is loaded, the browser creates a *Document Object Model* of the page. It's a tree object of HTML. HTML DOM is a W3C standard. There are also XML DOM standard similar to HTML DOM.

What should I know about it?

There is a '*document*' object in javascript, containing all serious methods and properties.

```
document.getElementById("myid").innerHTML("write something");
```



```
<div id="myid"></div>
```

```
<div id="myid">write something</div>
```

Javascript introduction


What is an event handler?

Usually when we are talking about events, we are talking about DOM events. For instance if something happens, an event handler will be triggered, which can be javascript function.


If user clicks an HTML element, it run 'onclick' event handler corresponding to this element.

What should I know about it?

In Object perspective event handlers are just callable functions. Browser will call these functions, if something happens in browser.



```
document.getElementById("myid").onclick=function(){/* do something */};
```



```
<div id="myid">click here</div>
```

```
<div id="myid" onclick="/* do something */">click here</div>
```


jQuery introduction

What is jQuery?

jQuery is a javascript library, designed to simplify client-side scripting of HTML.

What should I know about it?

There is a famous `\$` function, which makes everything easier.

Javascript: `document.getElementById("myid")`

jQuery: `$("#myid")`

Javascript: `document.getElementsByClassName("myclass").forEach(function(el){ el.onclick = function() { ... } });`

jQuery: `$(".myclass").click(function(){ ... });`

Javascript vs. Python

what is an object?

Javascript:

```
function MyClass(arg) {  
    var $this = this;  
    this.name = arg;  
    this.getName = function() {  
        return $this.name  
    };  
};  
  
var obj = new MyClass("my name");  
name = obj.getName();
```

Python:

```
class MyClass():  
    def __init__(arg):  
        self.name = arg  
  
    def get_name(self):  
        return self.name  
  
obj = MyClass("my name")  
name = obj.get_name()
```

Web Speech API

- It is not part of the HTML DOM. There are some voice activated elements in DOM but we are talking about another set of predefined objects:

- SpeechRecognition

SpeechRecognitionError
SpeechRecognitionAlternative
SpeechRecognitionResult
SpeechRecognitionList
SpeechRecognitionEvent
SpeechGrammar

- SpeechSynthesis

SpeechSynthesisUtterance
SpeechSynthesisEvent
SpeechSynthesisVoice
SpeechSynthesisVoiceList

Web Speech API code

speech synthesis

```
var msg = new SpeechSynthesisUtterance("say something");  
msg.onend = function() {  
    // do something;  
}  
speechSynthesis.speak(msg);
```

voice recognition

```
var recognition = new SpeechRecognition();  
recognition.onresult = function(event) {  
    if (event.results.length > 0) {  
        str = event.results[0][0].transcript;  
    }  
}  
recognition.start();
```

test it on browser's
developer console

Web Speech API

future and present

speech synthesis

- It will accept SSML. It's not implemented yet.
- The first implementation is on Chrome 33.

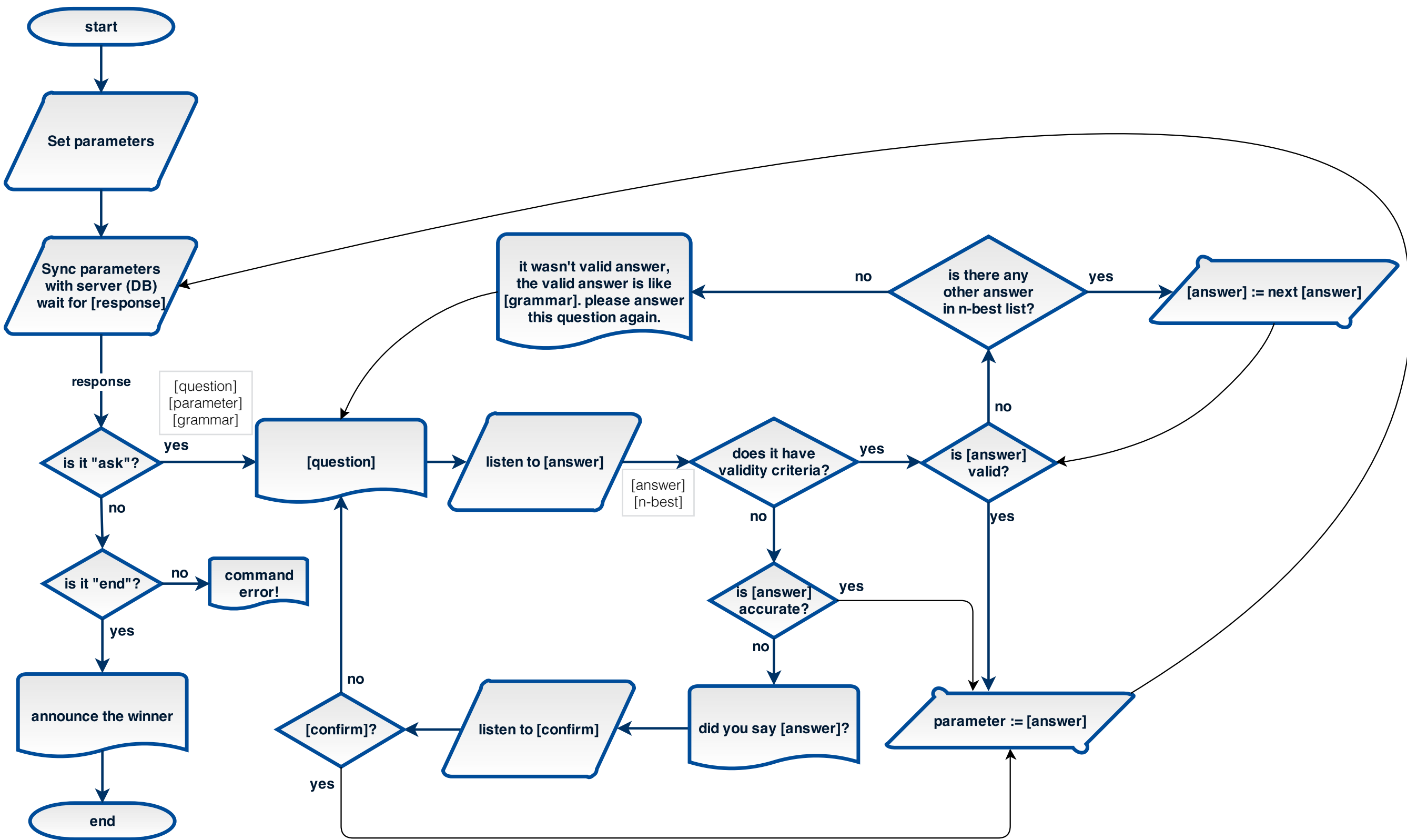
voice recognition

- It will accept SGRS. It's not implemented yet.
- Some parts still missing, for instance in n-best list, you don't have confidence rate.
- None of them are fully stable.
- None of them use the local OS speech engines yet.

Web Speech API

other issues

- With out grammar for voice recognition, recognizing all words are even harder.
In the next level, missing semantic part of grammar makes it even more complicated for the language understanding.
- There is a continues mode for voice recognition, but you cannot run it for ever. It is better to start listening when you are ready for it.
- The current speech synthesizer cannot read for ever, It just runs for limited time.



Server-side

- Dialogue manager
- Language understanding
- Language generation
- Database

Python as a web server

- There is a python library called SimpleHTTPServer
- The easiest CGI solution for Python is using it as a web server.
- It's not the best solution but it works fine with small projects.
- To even more simplification you can use Flask framework.

Flask introduction

What is Flask?

Flask is a microframework for Python to develop web application.

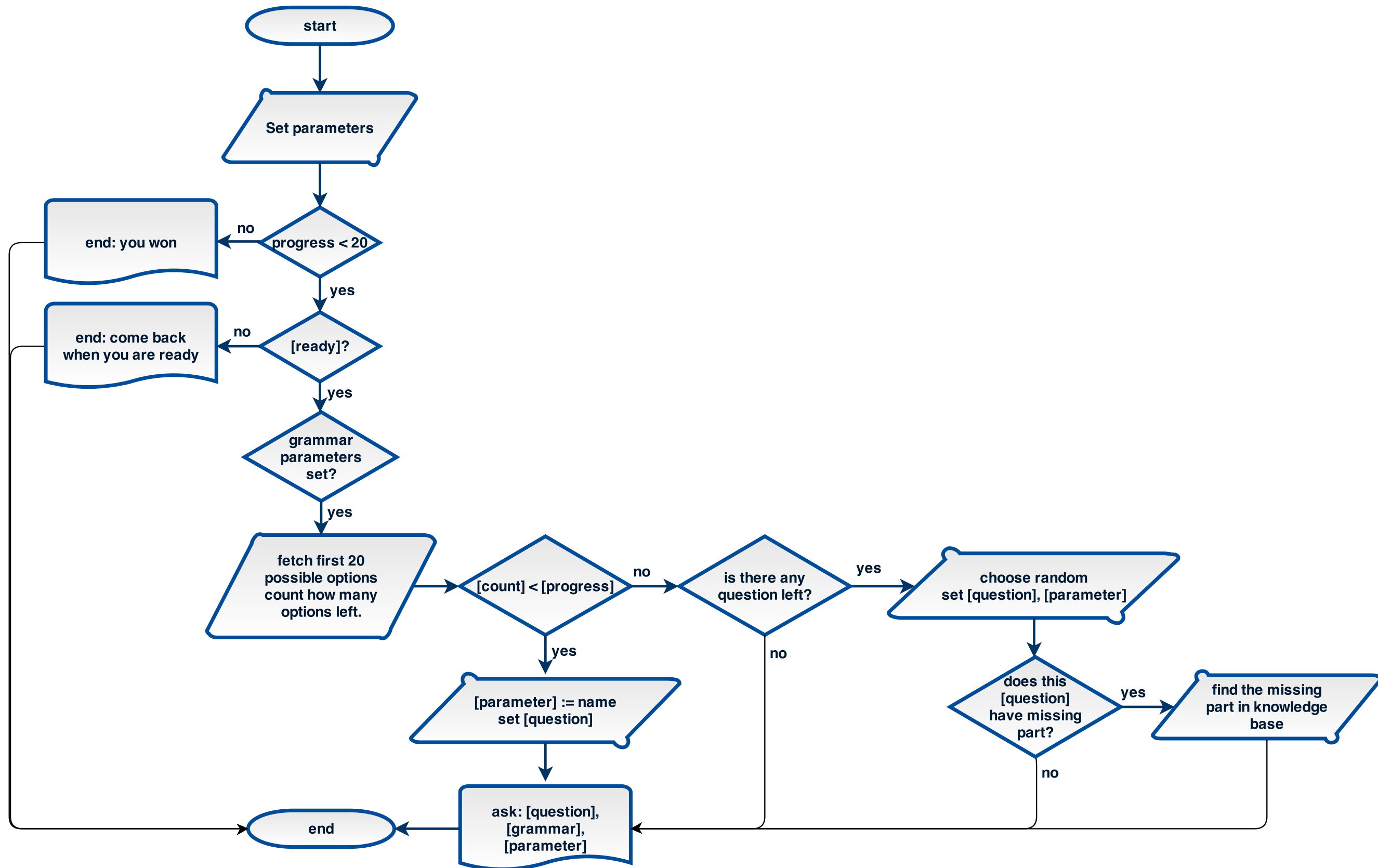
What should I know about it?

A function, with special wrapper `@app.route("/uri")` can be called with browser. Then you can pass arguments through `request` object.

```
from flask import Flask
app = Flask(__name__)

@app.route("/")
def hello():
    return "Hello World!"

if __name__ == "__main__":
    app.run()
```



Knowledge base

FreeBase introduction

- It's a collaborative knowledge base, consist of metadata.
(A semantic wikipedia.)
- It built by Metaweb company, and Google acquired Metaweb.
- They aimed to have ontologies like FOAF
(an ontology of persons)
- It has an open API. Sample python codes are available.
<https://developers.google.com/freebase/v1/search-overview>
- It is freely available for commercial and non-commercial use
under a Creative Commons Attribution License.

FreeBase API

a google service

```
{'filter': '(all type:/people/person /people/person/nationality:"Europ" /people/person/gender:"male")',  
'indent': 'true',  
'key': 'AlzaSyCNi0NvH5964WWfS80XrRmn9LzJCpTpWdg',  
'output': '(/people/person/nationality /people/person/profession /people/place_lived/location)',  
'query': ''}
```

https://www.googleapis.com/freebase/v1/search?filter=%28all+type%3A%2Fpeople%2Fperson+%2Fpeople%2Fperson%2Fnationality%3A%22Eurasia%22+%2Fpeople%2Fperson%2Fgender%3A%22male%22%29&query=&indent=true&key=AlzaSyCNi0NvH5964WWfS80XrRmn9LzJCpTpWdg&output=%28%2Fpeople%2Fperson%2Fnationality+%2Fpeople%2Fperson%2Fprofession+%2Fpeople%2Fplace_lived%2Flocation%29

```
{  
  "status": "200 OK",  
  "result": [  
    {  
      "mid": "/m/082db",  
      "id": "/en/amadeus_mozart_wolfgang",  
      "name": "Wolfgang Amadeus Mozart",  
      "notable": {  
        "name": "Opera Artist",  
        "id": "/m/05lls"  
      },  
      ....  
    },  
    ...  
  ]  
}
```

Play with it

Future developments

- Extend the categories to other fields, other than people.
- Empower it with an artificial intelligence to have a reasoning unit and decision making unit, instead of randomly pick a questions.
- Add more capabilities over interpretation of data.
- ...