

# How the Matchmaking Works



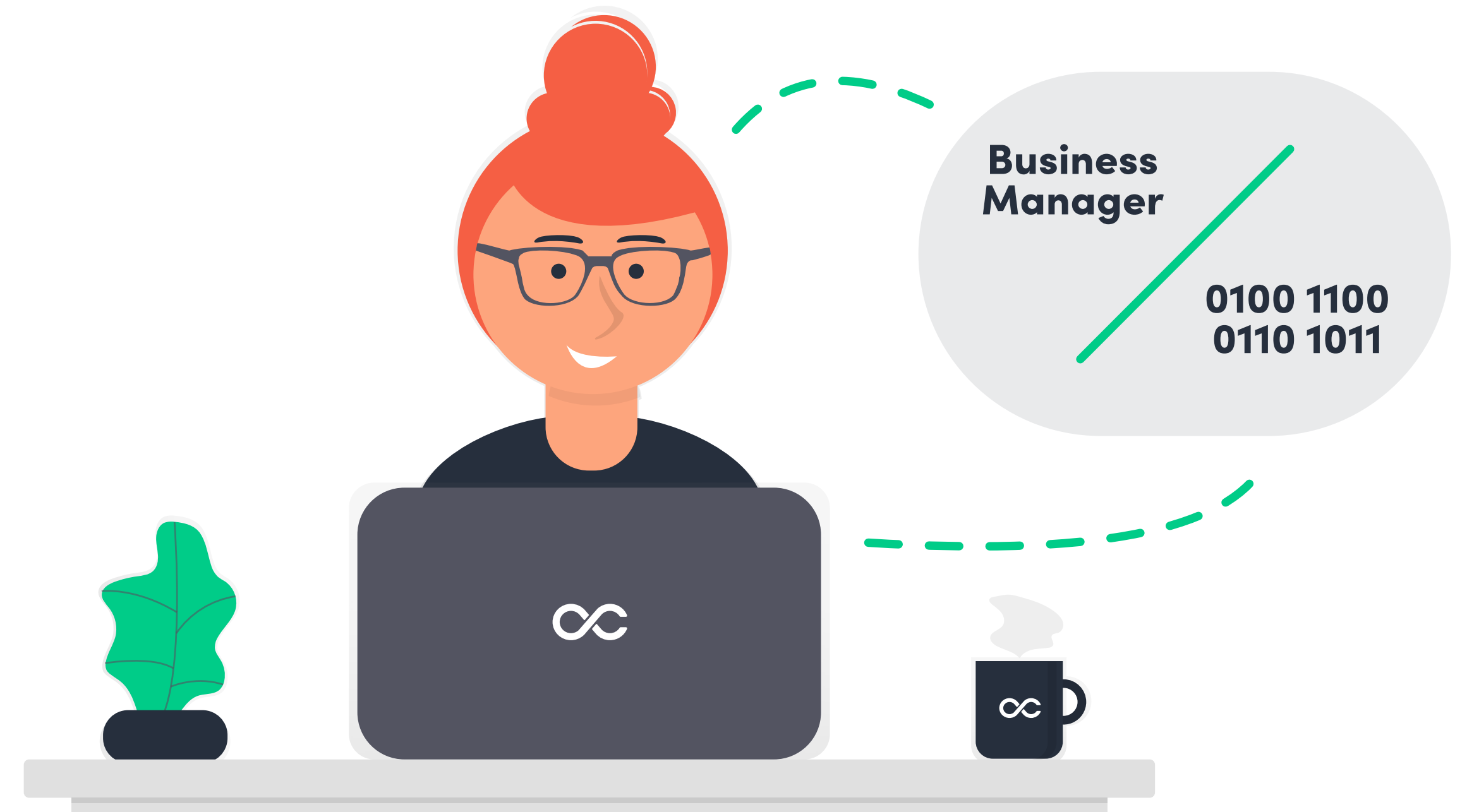
swapcard

Empower Meaningful Encounters

# The NLP Pipeline

NLP stands for **Natural Language Processing**. It is the techniques that allow computers to process and understand human language.

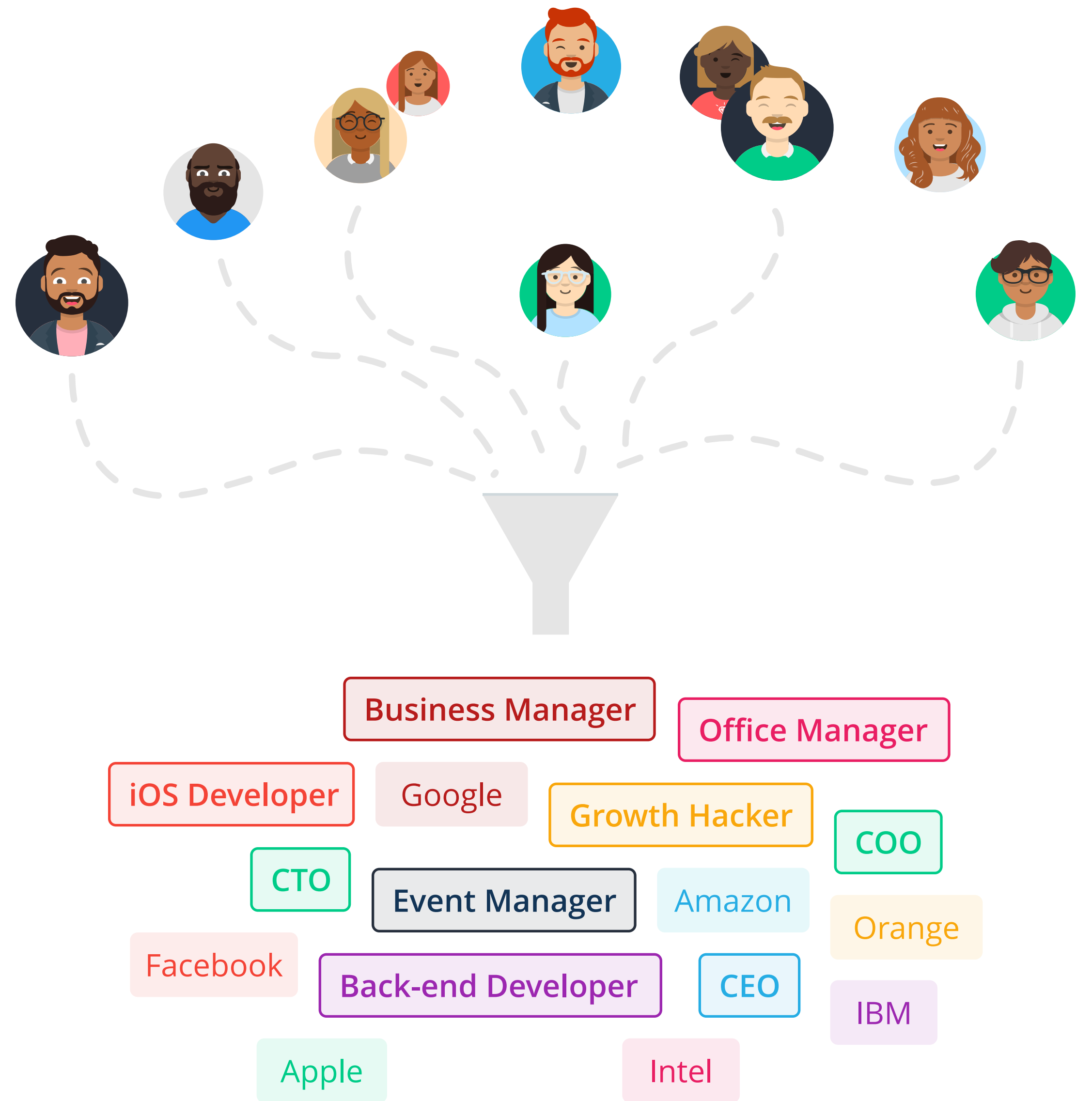
At Swapcard we use NLP to filter important **keywords** out of the data provided to us by the attendees. To do this we use different technologies: *SpaCy*, *Word2Vec*, and *ElasticSearch*.



# The NLP Pipeline

The NLP pipeline is comprised of different parts:

- The listener that constantly monitors changes made by users to their profiles.
- The NLP pipeline, which cleans the data, checks it against databases, and extracts the meaningful keywords such as skills or companies.



# The NLP Pipeline

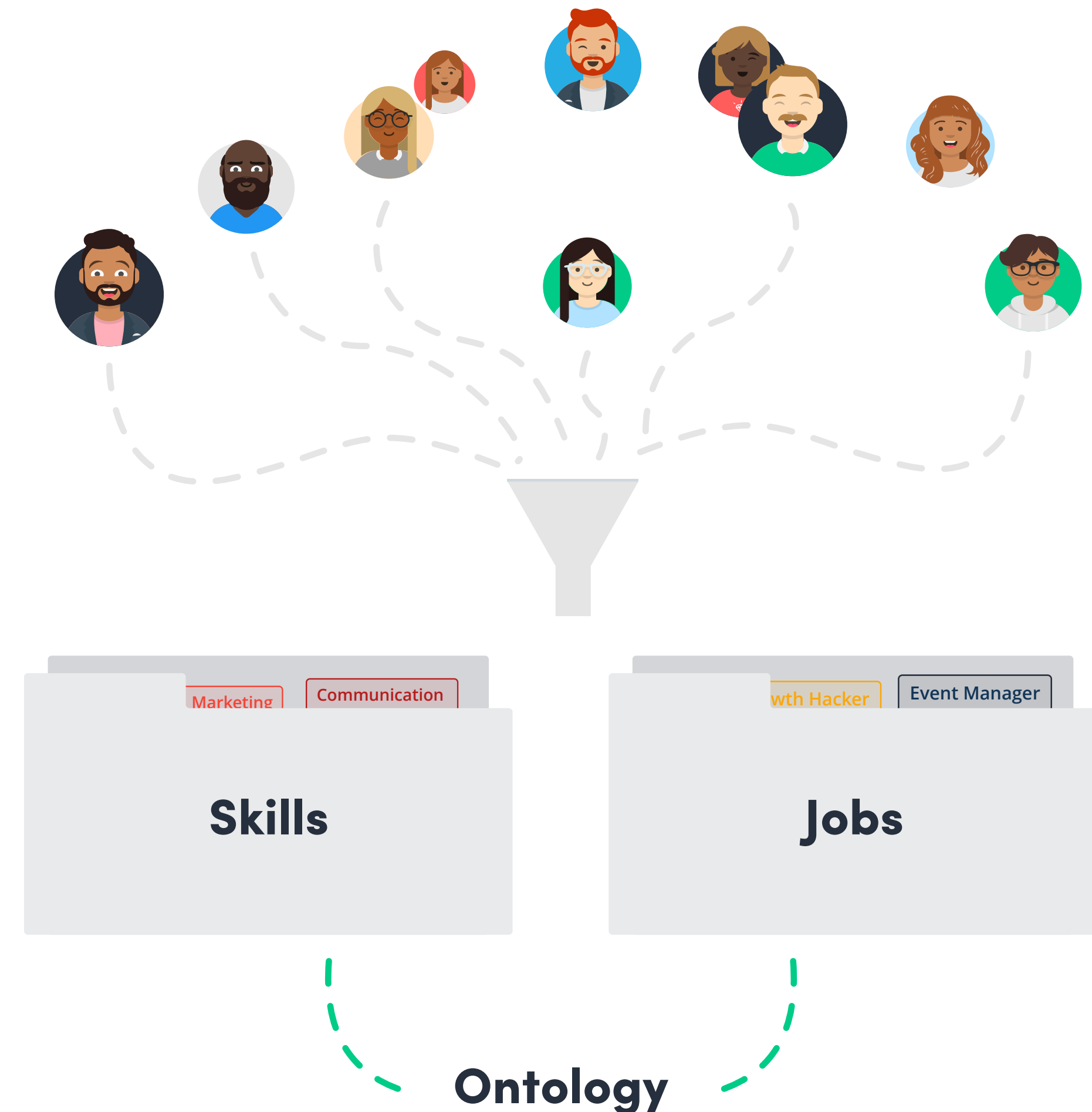
To do this we use data provided by users as well as machine learning techniques :

- *Named Entity Recognition* and *POS Tagging* to recognize the part of the sentence that are **jobs** and **skills**
- *Dependency Parsing* to find which words depend on which.

For example in the job title '*C++ Ninja*' we use *Dependency Parsing* to find which is the **meaningful word** (C++).

We then send all that data about users to our databases and to ElasticSearch, which provides near **instant responses** to queries about our users.

When the NLP pipeline has done its job, the collected data is used to create the ontology.



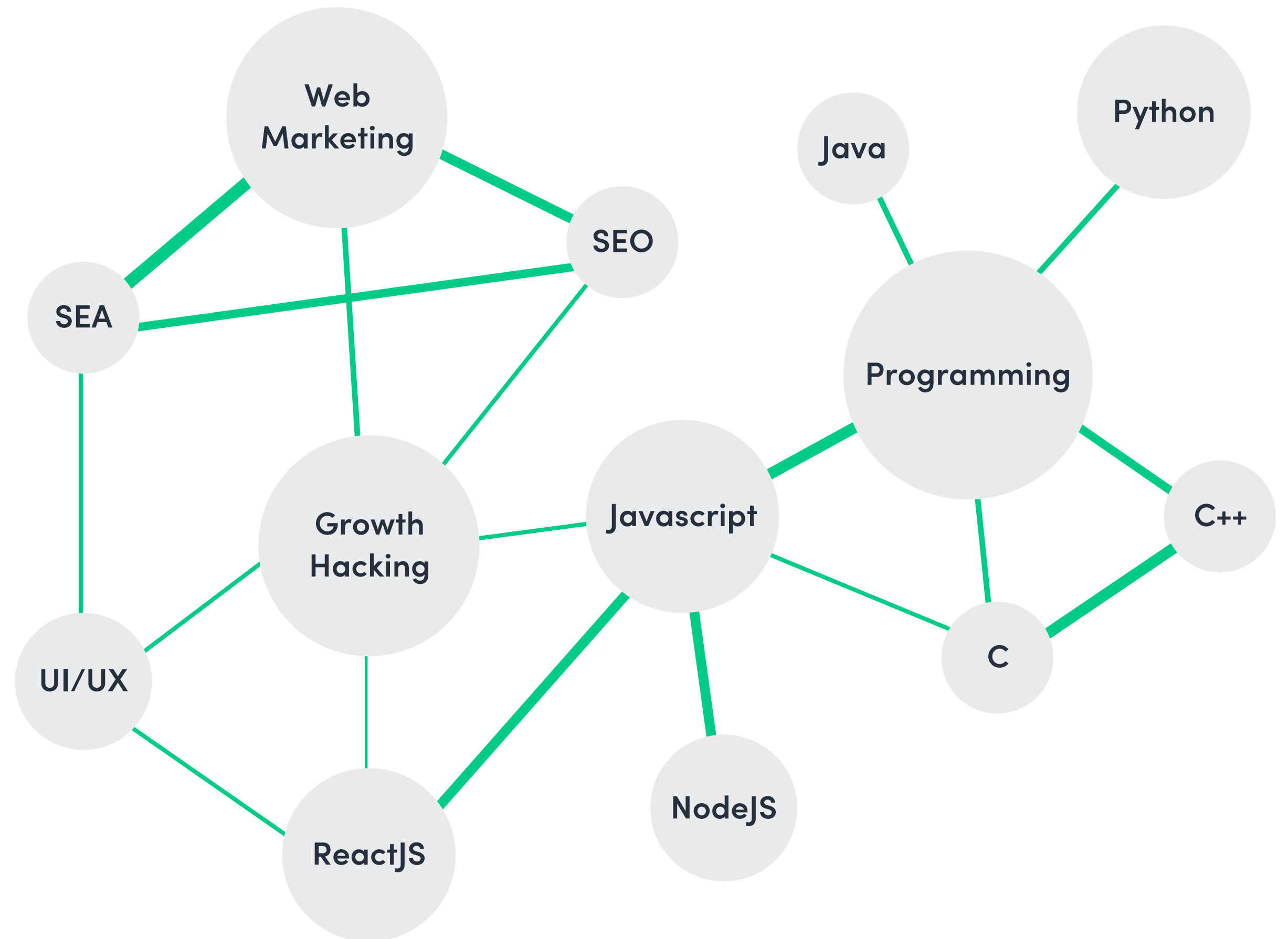
# The Ontology

An ontology is a way to describe the world as a **graph** (nodes connected by links) in a meaningful way. This allows us to represent all the data we collect as a coherent set.

Within Swapcard our ontology has **nodes** that correspond to **characteristics, keywords, companies** and **industry** sectors. When an item has multiple keywords, these keywords are linked together.

The links represent the **strength** of a connection between a pair of keyword is linked. The more often two keywords appear together, the stronger the link.

With enough data we can then create **groups** of keywords that belong together because they are strongly linked together.



# The Ontology

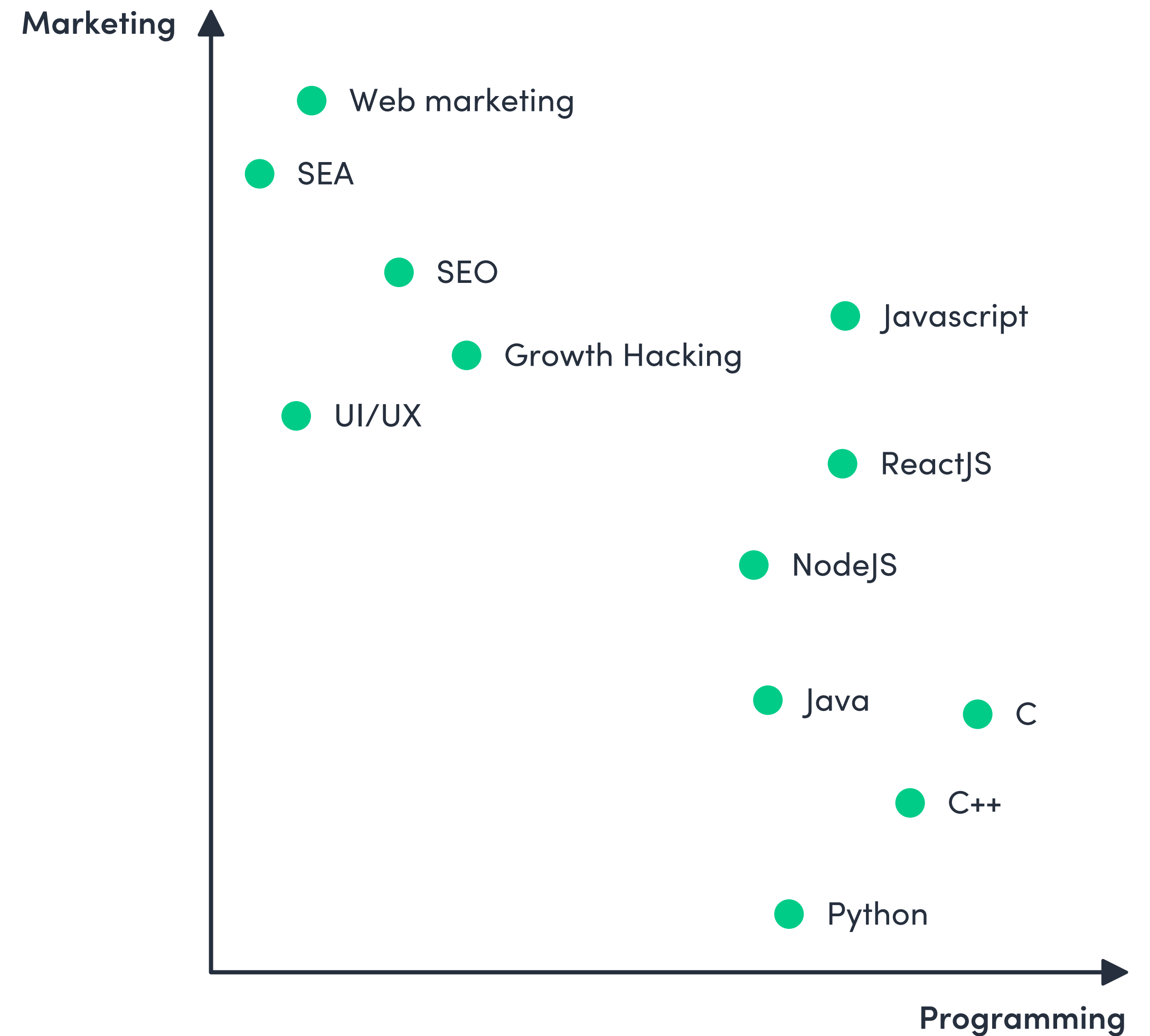
Each node is then embedded into a high dimensional vector, all **within the same space**.

This space contains all the **relevant** information about our data. Each item is a single point in this space.

If the distance between two points is short, it means that these two items are **closely related** and that they were close together in the ontology.

These vectors are at the **heart** of our matchmaking system. They are useful because machine learning often requires vector data.

Once the ontology has been established we can move forward to the **matching system**.

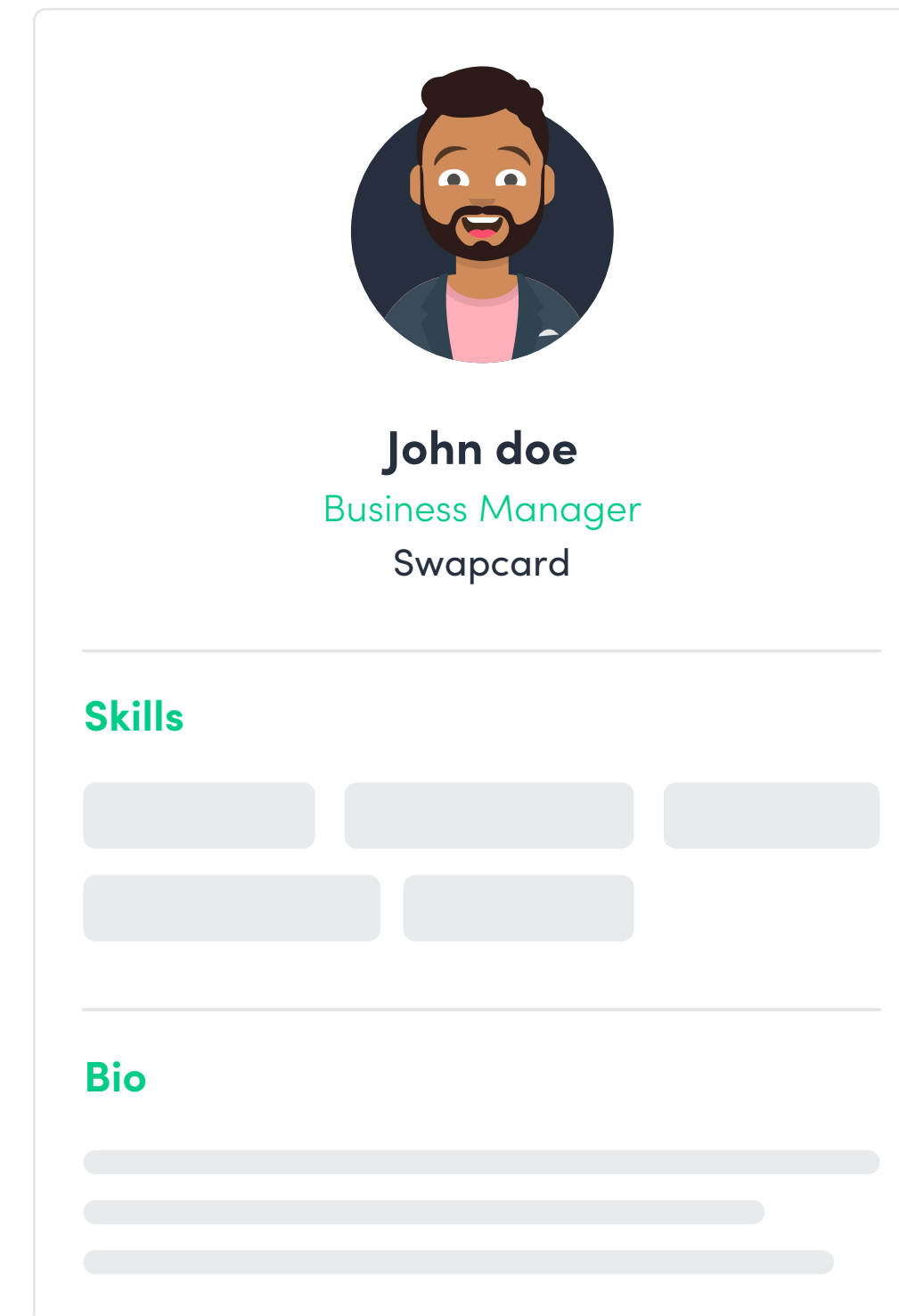


# The matching system

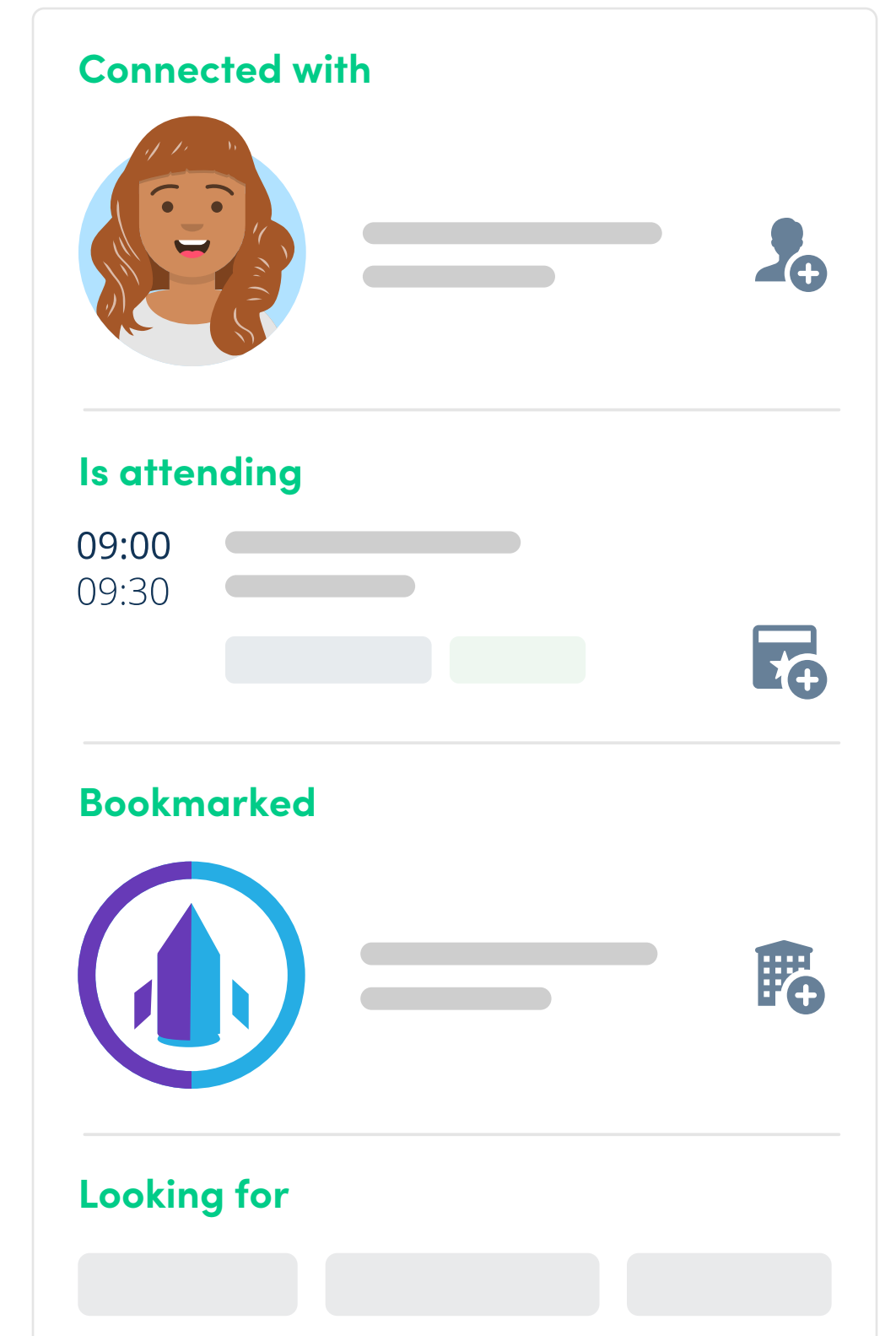
All items (users, sessions, products and exhibitors) are represented by a singular **profile**, based on their data such as **keywords, skills or industries**.

Users also have a second profile to represent their **activity** within the app : Who they **connected** with, which **session** they went to, which **exhibitors** were bookmarked, what they are **looking for**.

**First profile**  
to represent them



**Second profile**  
to represent their activity







John doe

# The matching system

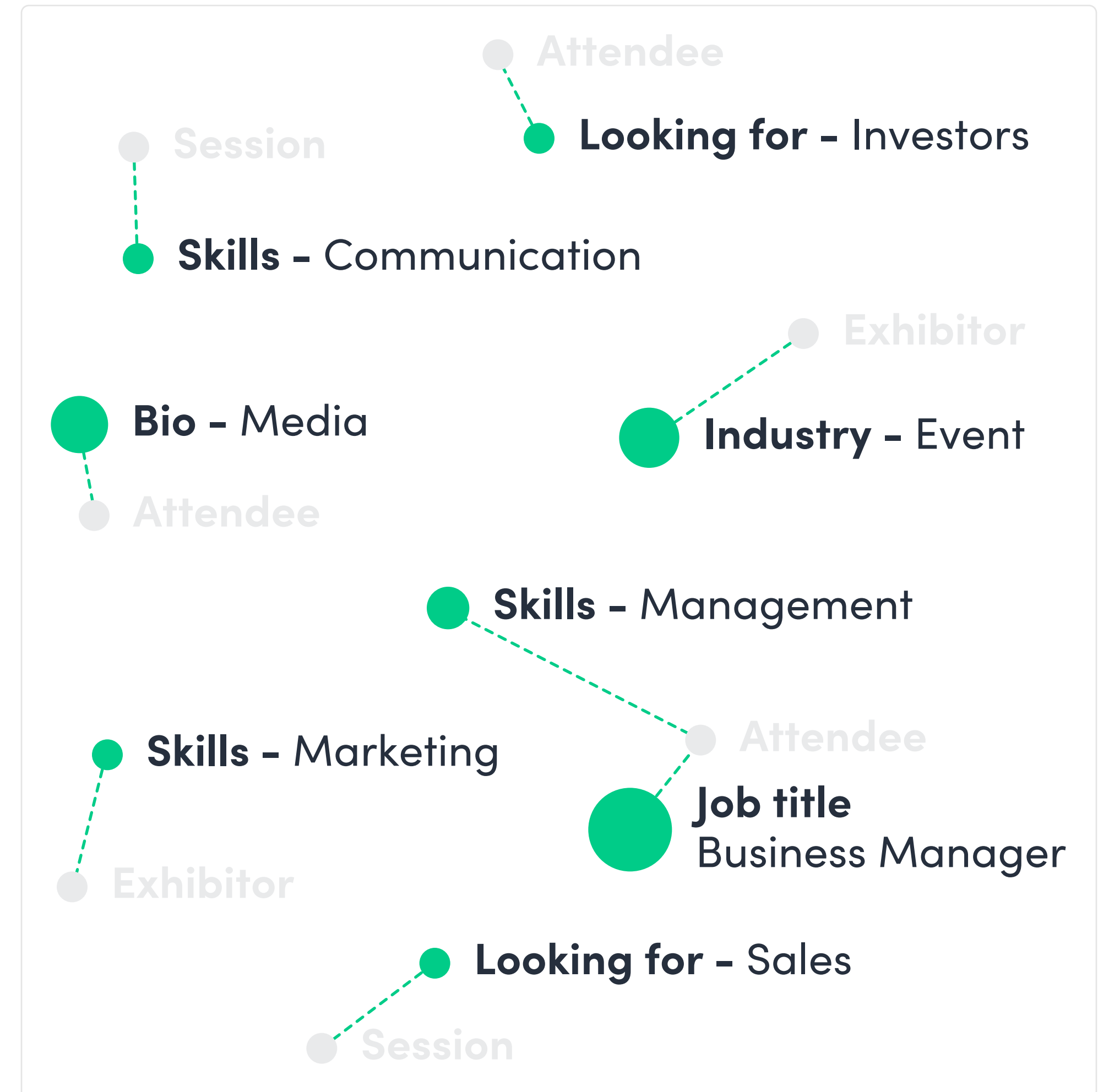
Profiles are mathematical representations that are generated from different **keywords**.





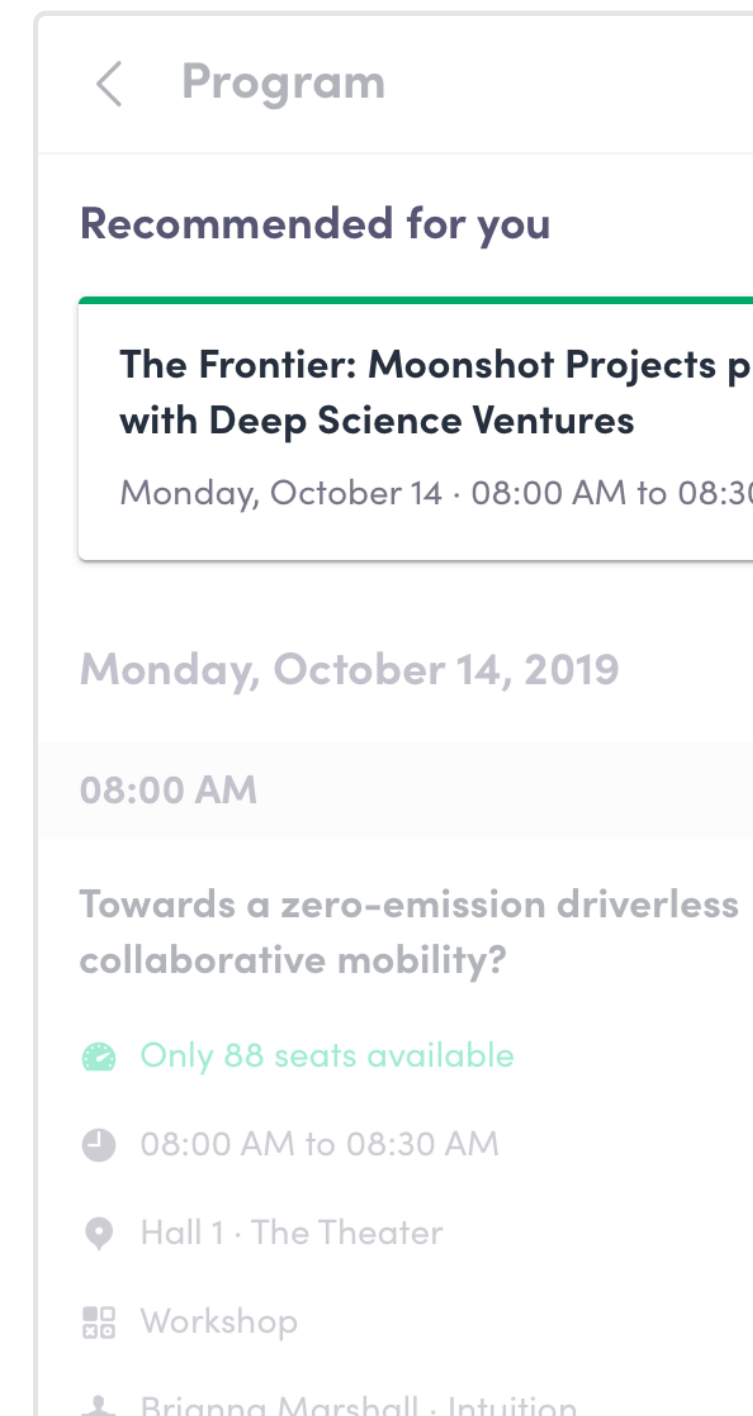
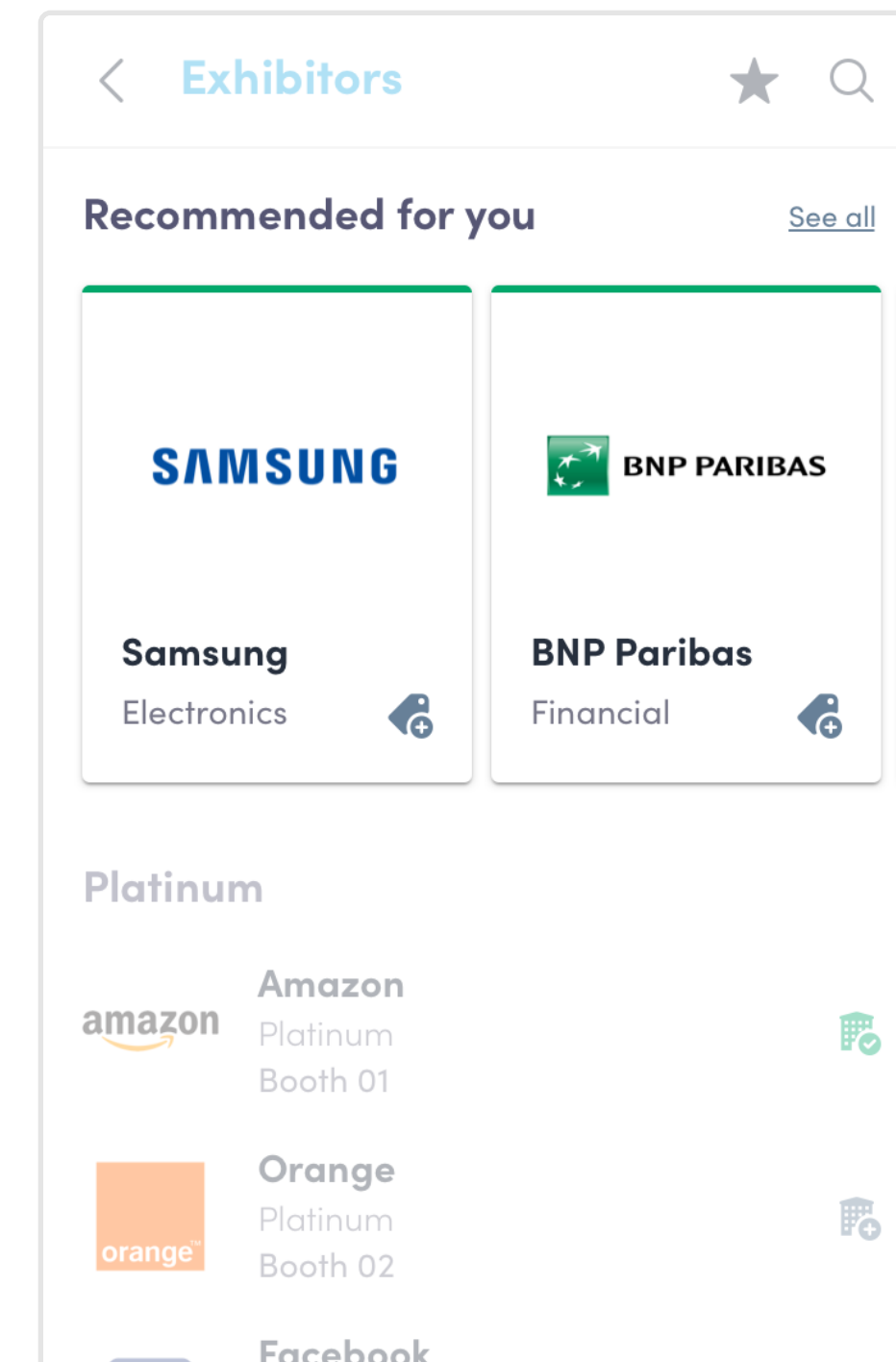
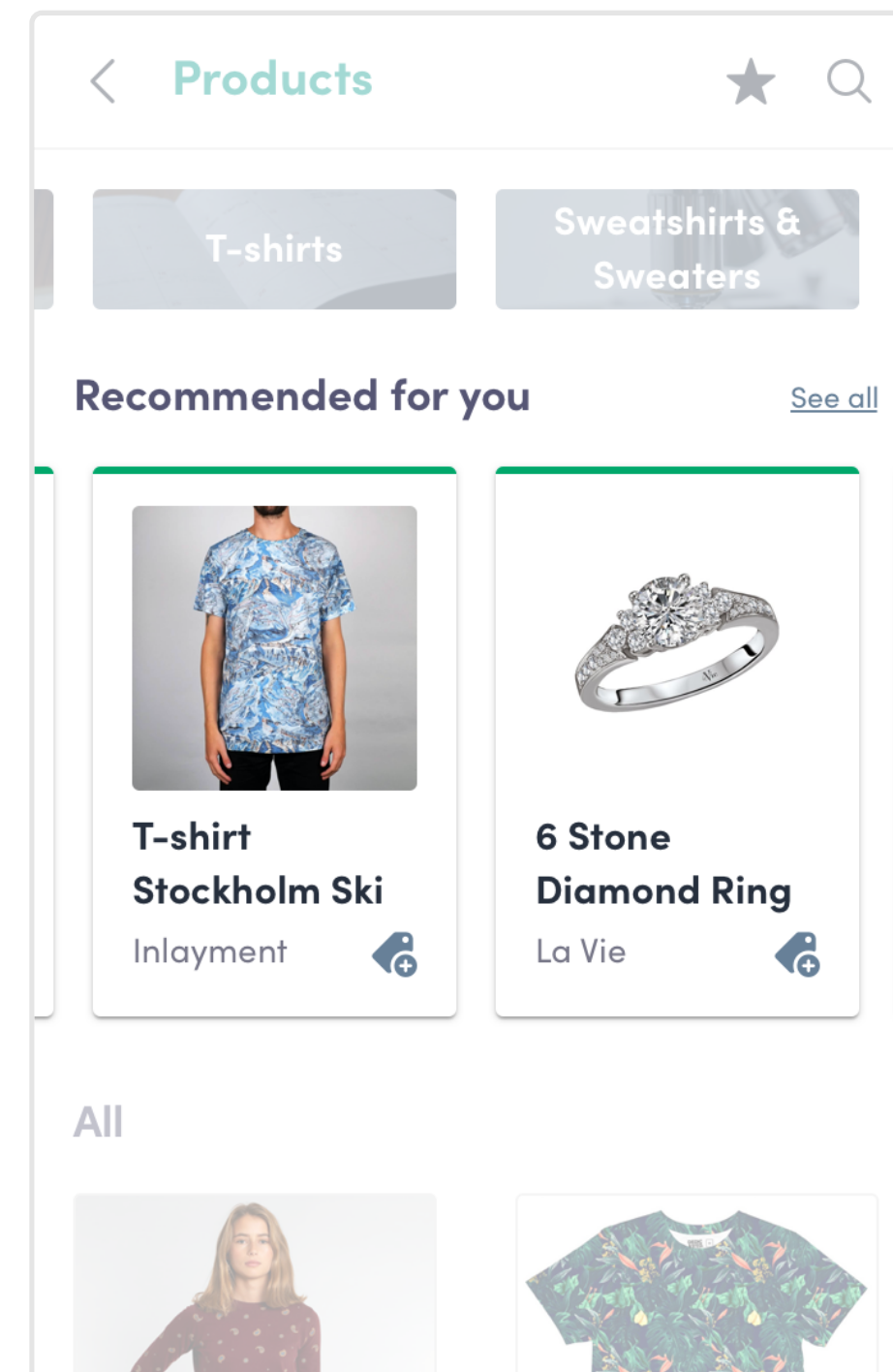
# The matching system

We calculate **distance** between what users are looking for and the **items available** at the event (sessions, exhibitors, products, other attendees) and use them to generate our **recommendations**.



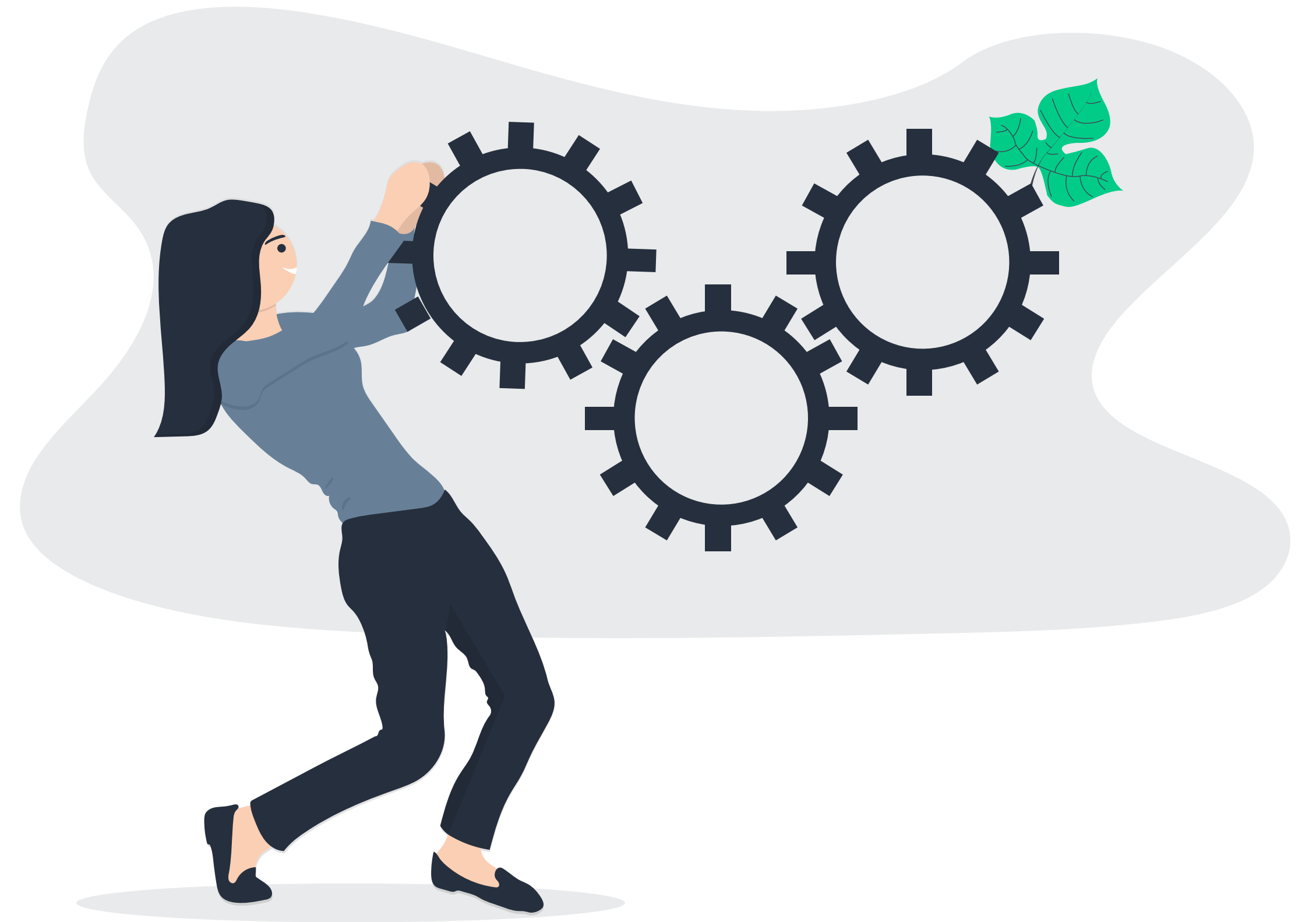
# Suggestions

We also use this technique to generate a list of similar items that will allow us to propose other items of interest when a user is looking at a specific session, product or exhibitor.



# Feedback

The NLP pipeline and Ontology steps feed into each other to **improve** the ontology over time. When new links are created between nodes because of user input and NLP updates, the ontology is **updated**. When the ontology is updated, new associations between keywords can be **created** for users through the NLP pipeline.



**Thank you!**

**swapcard**

Empower Meaningful Encounters