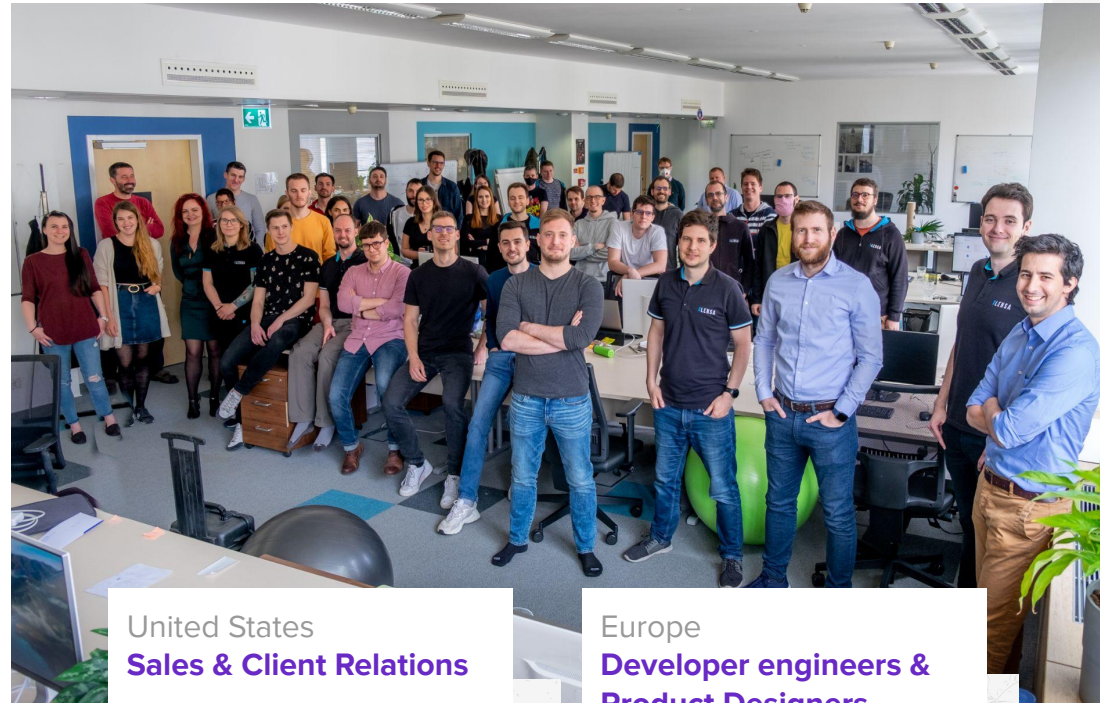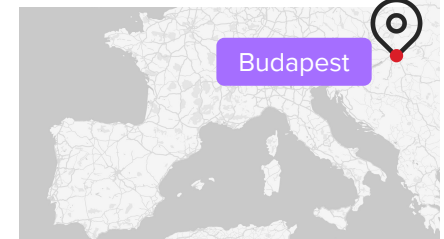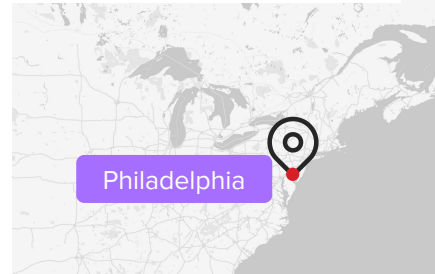# LENSA

Who Are We?

A career site that helps job seekers with career pathing, informational resources, job search tools and other career-related services.

United States
Sales & Client Relations

Europe
Developer engineers & Product Designers

Philadelphia

Budapest

Let's start with some large numbers

Job listings

160 million
Job postings

4.5 million
Clicks generated monthly on job content

Members

17+ million
Jobseeker members

475K+
new users monthly

Daily notifications
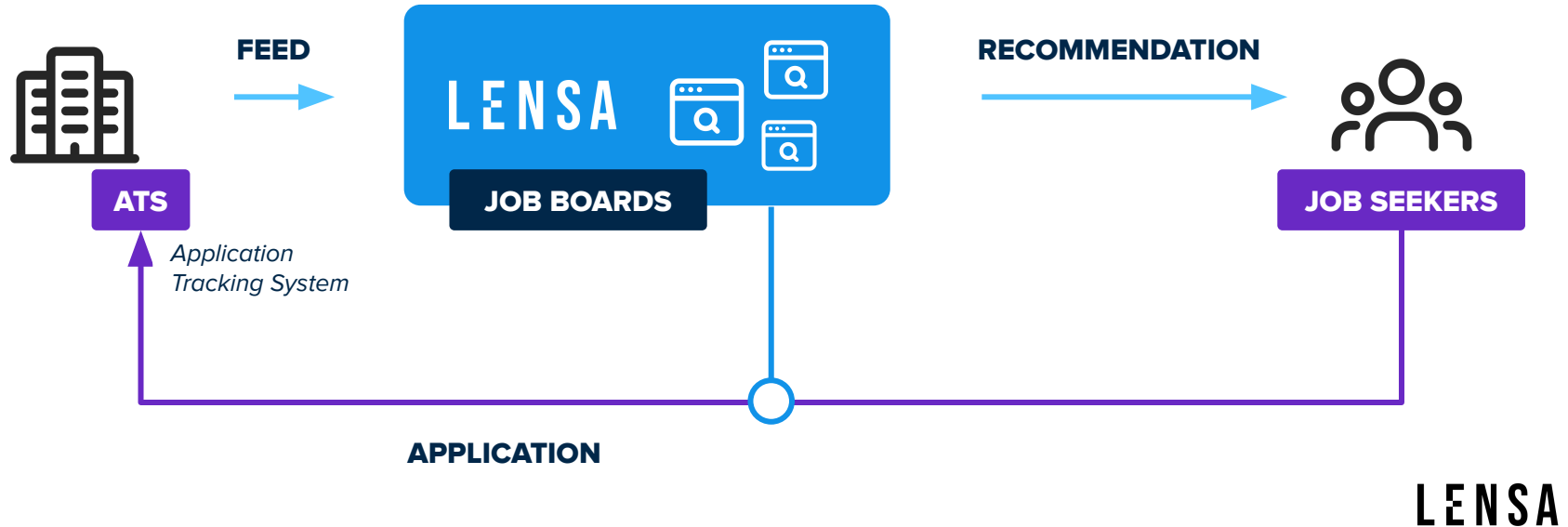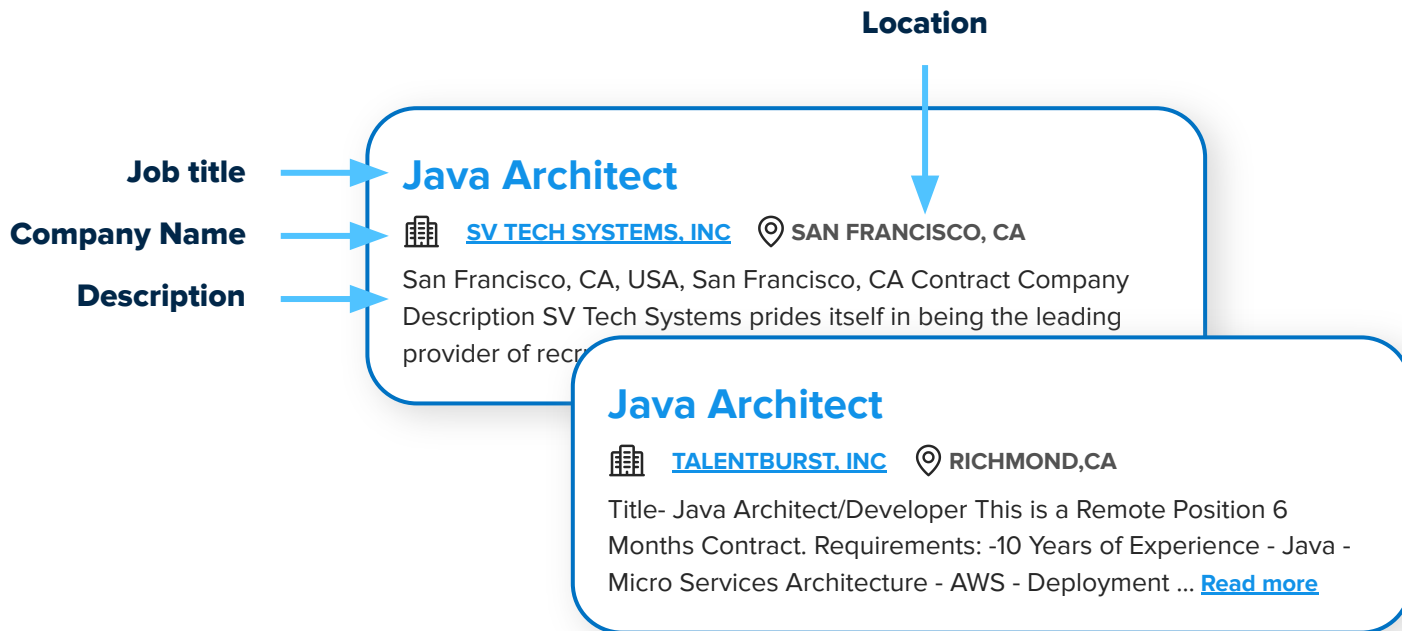
6.6 million
nightly emails sent daily

500K
SMS sent daily

LENSA

How the online job market works in the US?

FEED

LENSA

JOB BOARDS

RECOMMENDATION

ATS

Application Tracking System

JOB SEEKERS

APPLICATION

LENSA

How does a job posting look like?

Location

Job title → **Java Architect**

Company Name → 🏢 **SV TECH SYSTEMS, INC** 📍 **SAN FRANCISCO, CA**

Description → San Francisco, CA, USA, San Francisco, CA Contract Company Description SV Tech Systems prides itself in being the leading provider of recr

Java Architect

🏢 **TALENTBURST, INC** 📍 **RICHMOND,CA**

Title- Java Architect/Developer This is a Remote Position 6 Months Contract. Requirements: -10 Years of Experience - Java - Micro Services Architecture - AWS - Deployment ... **Read more**

LENSA

Internet is messy

Welcome to Cajun Steamer Bar & Grill!

TALENTBURST, INC   RICHMOND,CA

Who We Are

If you've got that craving for authentic Cajun cuisine, then you've come to the right place! At Cajun Steamer ... **Read more**

<h3>Why a Great Opportunity</h3><p>

TALENTBURST, INC   RICHMOND,CA

This is with one of the premier providers of consumer products&nbsp;in the world!&nbsp; Terrific work environment and tons of room for career growth.&nbsp; ... **Read more**

LENSA

Why we need a job title predictor?

Main problem

- Employer's HR wants to write catchy job titles

- Our job processing/BI wants to categorise jobs as clearly as possible

Raw job title	Cleaned job title
$1000-$2000 sign on bonus! now hiring warehouse associates	warehouse associate
area sales representative (must live in greater richmond area)	sales representative
pmhnp nurse practitioner	Psychiatric Mental-Health Nurse Practitioner
pahrtners deaf services	?

LENSA

What is a job title?

- **Raw job titles can be anything**
  Quasi infinite number of possible inputs

- **Try to keep coverage high**
  We can't just throw out non-standard job titles

- **Cleaned job titles come from a relatively small set**

- **Our data is biased towards our jobs**

LENSA

What is a job title?

Golf Lawn Maintainer

Gardener Apprentice

Bush Whacker

Gardener

We don't care about modifiers

- Like senior/junior/associate

- These modifiers are extracted by a separate model

- Although Associate can be a job title on its own

LENSA

Collect title set

ONET Online   Wikipedia   WWW

MERGE / DEDUP

18k
curated job title

LENSA

Normalization (overview)

18k clean — **the rest, raw**

| Truck Driver | Gardener | Data Scientist | Police Officer | Bowling ball cleaner | Gardener at a golf court | Come and find your dream job for $38 per hour | Oyster Shucker | Law Enforcement Officer |

Truck Driver: Duties include: Assisting and supporting the implementation of EHS initiatives and strategies to reduce risk including but not limited to …

Gardener: Do you love working in your garden? Garden professionally and start earning money now! We're REAL gardeners, not landscapers…

Data Scientist: At Lensa we're looking for a talented, hardworking passionate data scientist who can help us build sophisticated machine learning models that will…

Police Officer: The City of Leawood offers career development opportunities in which the employee could increase their experience, training, and pay.

Bowling ball cleaner: SONIC Drive-In - 1392 E. FLORENCE BLVD. [Food Service Worker / Wait Staff] As a Carhop or Skating Carhop at Sonic, you'll: Ensure that all restaurant menu items meet operational…

Gardener at a golf court: Responsibilities: Trim hedges and plants to proper size and shape Mow and edge lawn areas to maintain neat appearance Apply correct amount of water…

Come and find your dream job for $38 per hour: Earn up to $38 per hour driving with Lyft in Phoenix. New drivers can earn $3,500 in Phoenix after 180 rides in 30 days. Terms Apply What is Lyft? Lyft is a platform that…

Oyster Shucker: We are seeking a Oyster Shucker with upscale, high-volume restaurant experience. Working closely with nationally-acclaimed culinary director…
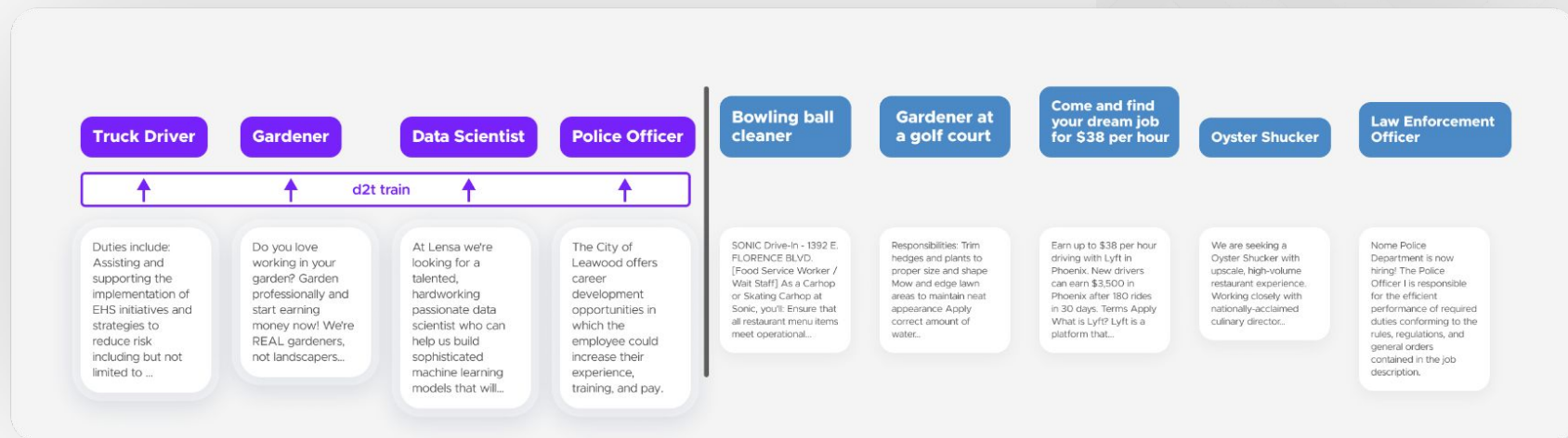
Law Enforcement Officer: Nome Police Department is now hiring! The Police Officer I is responsible for the efficient performance of required duties conforming to the rules, regulations, and general orders contained in the job description.

■ Map from an open set to a well defined, finite subset

■ How to get training data?

　○ Learn from (any) coupled data!

LENSA

Normalization (overview)

- Map from an open set to a well defined, finite subset

- How to get training data?

  - Learn from (any) coupled data!
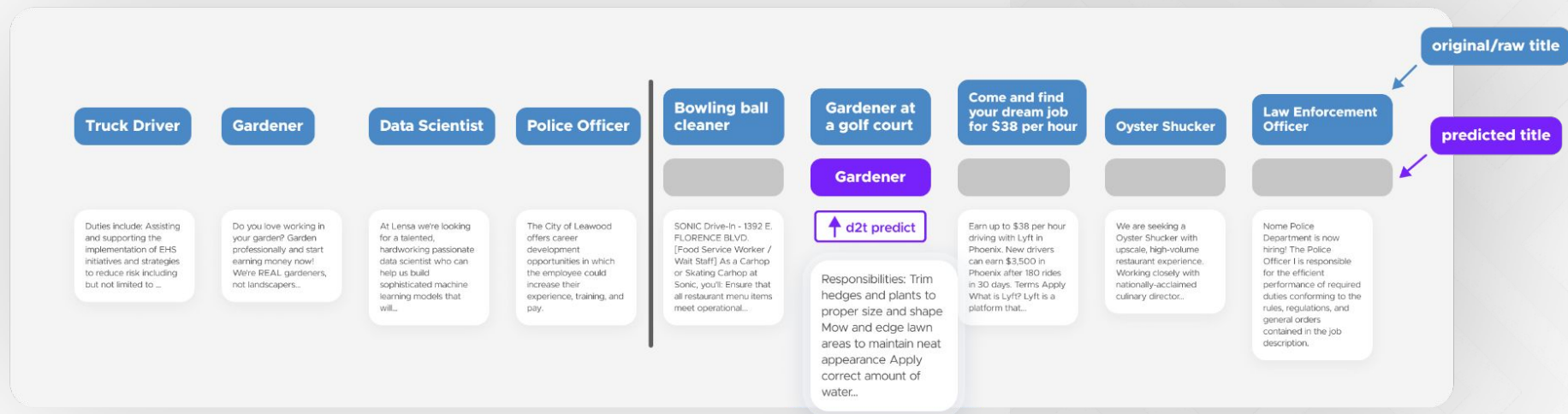
LENSA

Normalization (overview)

- Map from an open set to a well defined, finite subset
- How to get training data?
  - Learn from (any) coupled data!

Normalization (overview)

- Map from an open set to a well defined, finite subset

- How to get training data?

  - Learn from (any) coupled data!

Normalization (pseudo code)

- Obtain title+description pairs

- Select job descriptions where the job title exactly matches our predefined list

- Train **description2title** predictor (classifier) on those jobs

- Predict (cleaned) job title with previously trained predictor on the rest of the dataset

- Lose the descriptions

- Train a **title2title** predictor (classifier) on the generated data

LENSA

Normalization (specifics)

NN Architecture

- tokenization
- word embedding
- convolution layers
- softmax (throw out predictions that are not confident)

Data

- Augmented with trivial descriptions
- Down-sampled
- Up-sampled
- Try to balance it

Similar to *title2title* model

LENSA

Take away

If you have **a bunch of data** and a **subset to map to**
and **any coupled data with it**,
then **you can train a normalizer this way**

LENSA

Thank you, feel free to ask

Gábor Borbély
Data Scientist at Lensa.com
gabor.borbely@lensa.com

LENSA