VISUAL SEARCH & DATA SCIENCE ON STYRIA'S CLASSIFIEDS

Davor Anicic

Business Development and Sales Manager

Styria Digital Services

ONE SPIRIT,
UNLIMITED IDEAS

(STYRIA

THE REGIONAL PUBLISHER
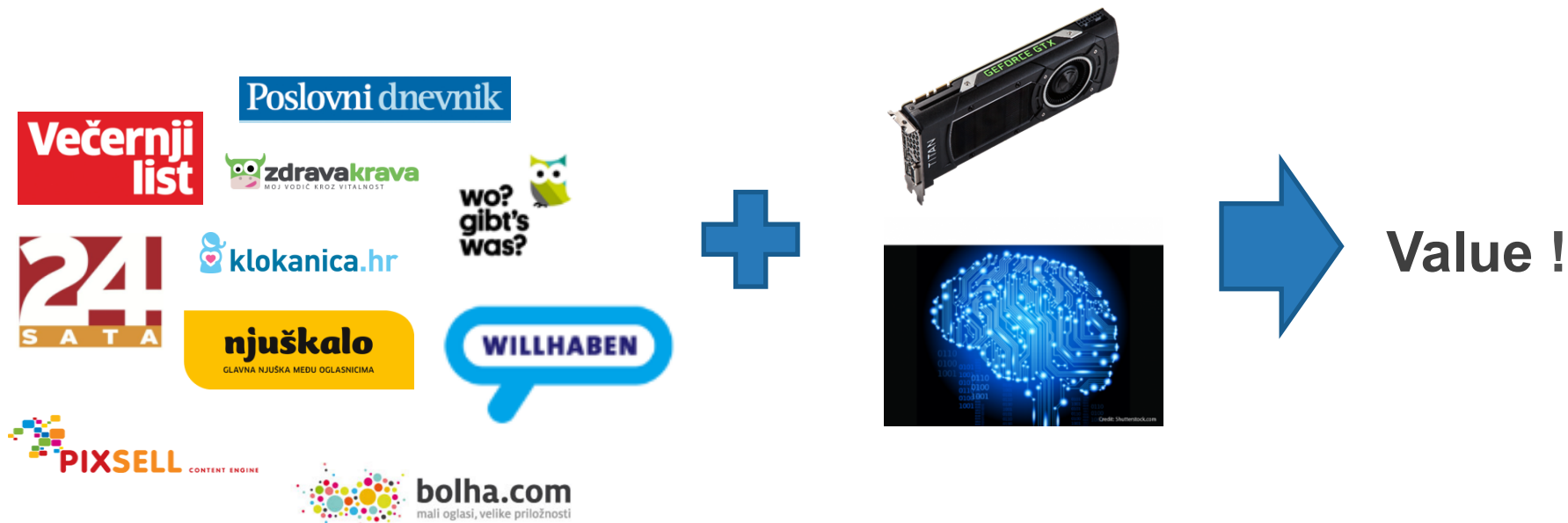
3 classifieds - market leaders

DATA SCIENCE TEAM

- 8 data scientists and programmers

- Initiative from 2014, team formed in 2015

- Machine learning experts:

  Computer Vision and Natural Language Processing

WHAT DO WE DO?

STYRIA

Data:
- text
- images
- users

Value !

- Computing resources
- Machine learning expertise

TECH. THAT MAKES THE DIFFERENCE

- Machine Learning (Deep Neural Networks)
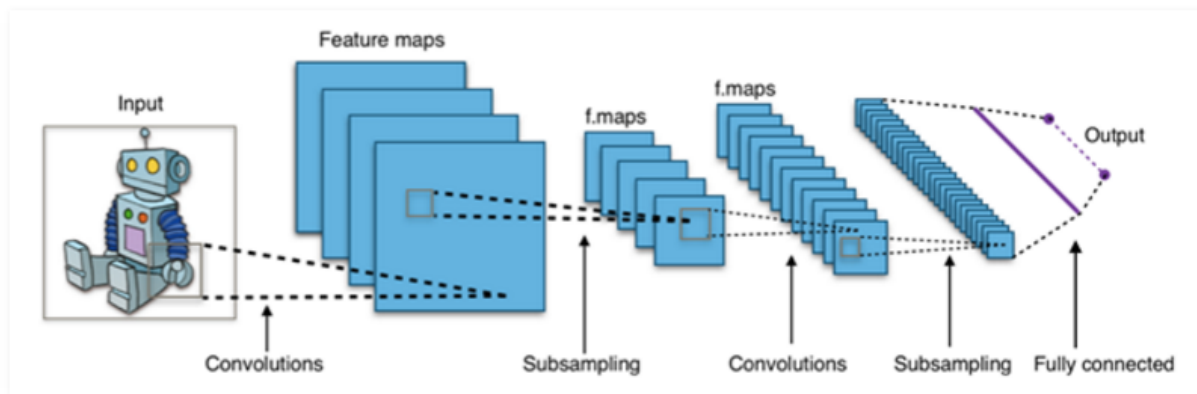- Program is not explicitly programmed, but the solution is learned from the data
- R&D

Figure 1: Typical architecture of a Convolutional Neural Network. Source: Aphex34 (Own work) CC BY-SA 4.0, from Wikimedia Commons
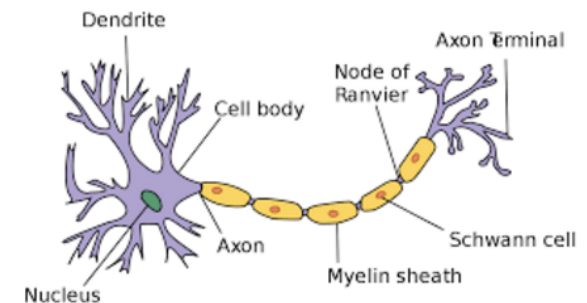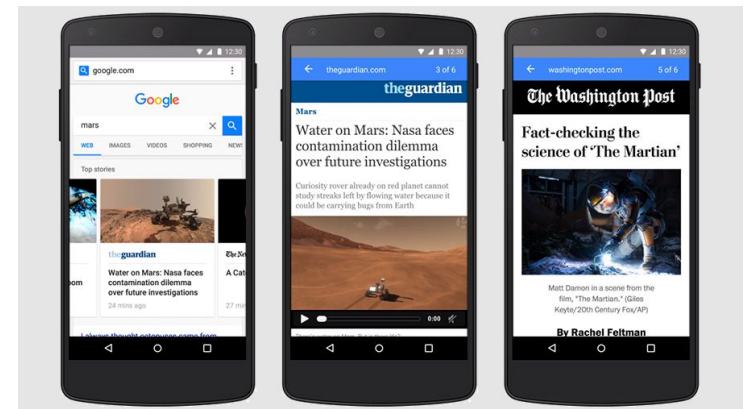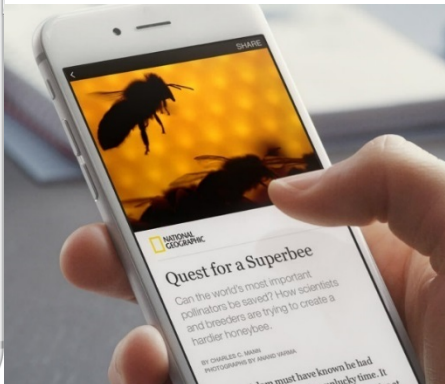
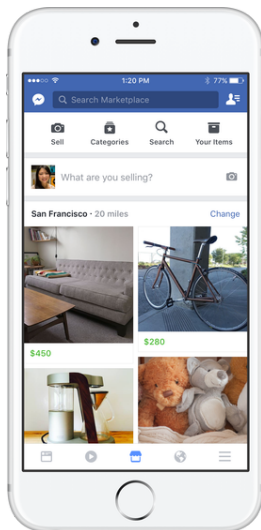Figure 2: Wikipedia

COMPETITION?

IS IT THE GAME FOR A PUBLISHER OR A CLASSIFIED?

- specific dataset and classification needs
- small and agile (internal) team
- understand your data

→ We can compete with the biggest players!

VALUE FOR THE CUSTOMER

| Seller's side	Buyer's side
Goal - easy & fast ad placement **How** - automatically categorise user's ad - take the photo → we do the rest	**Goal -** finding visually similar images **How** - recognise what is in the image - find similar

SPECIFICALLY DEVELOPED FOR CLASSIFIEDS

- Own data used for training the models
- Integrated to keep the data synced
- Visually–Semanticity tradeoff for fine-tuning
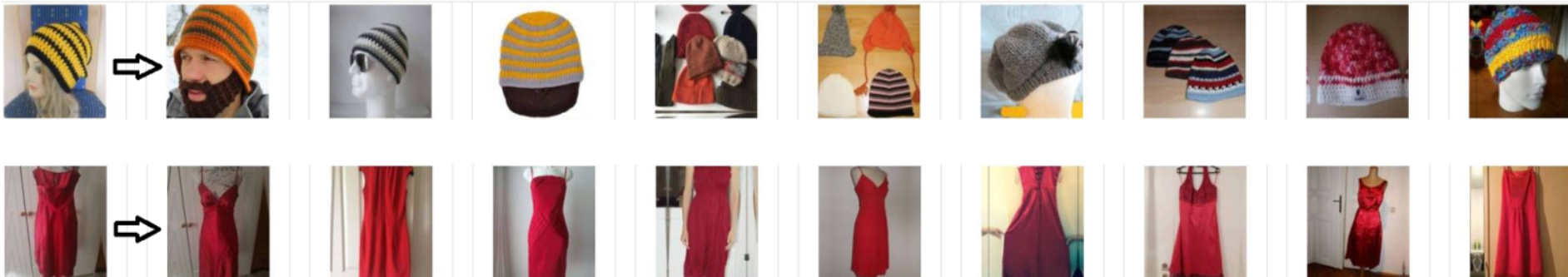
WILLHABEN APP (IOS & ANDROID) &
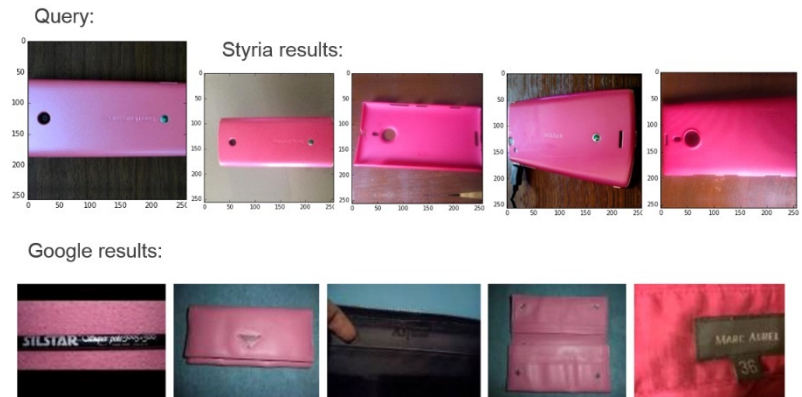## MOBILE.WILLHABEN.AT
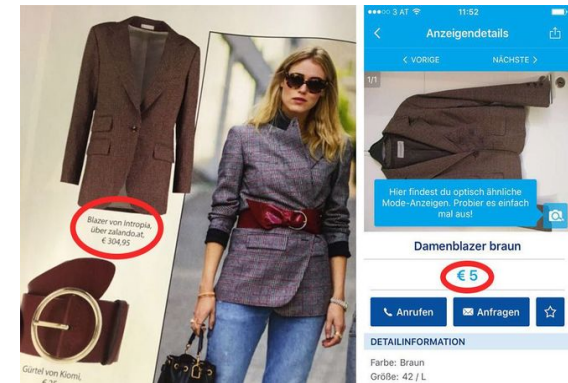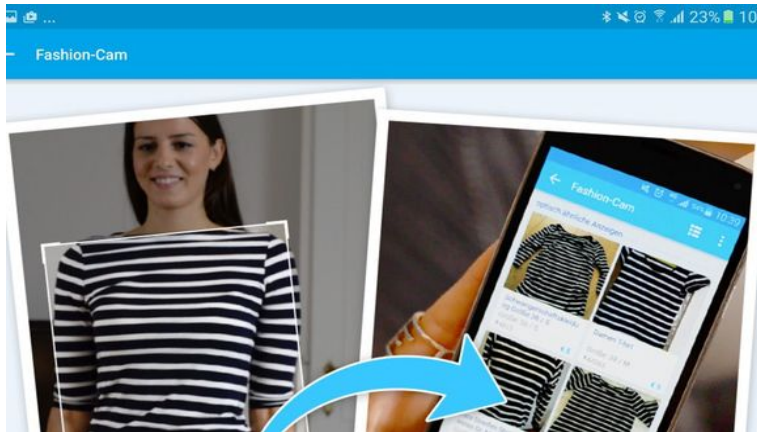
- 230 fashion categories
- 5 million images training set
- 3.2 million neurones with 6.1 million weights
- 200ms results returned to the user

Promo video: https://vimeo.com/191007189?from=outro-embed
Tech blog: http://tech.willhaben.at/2016/11/visual-search-applied-fashion-cam.html
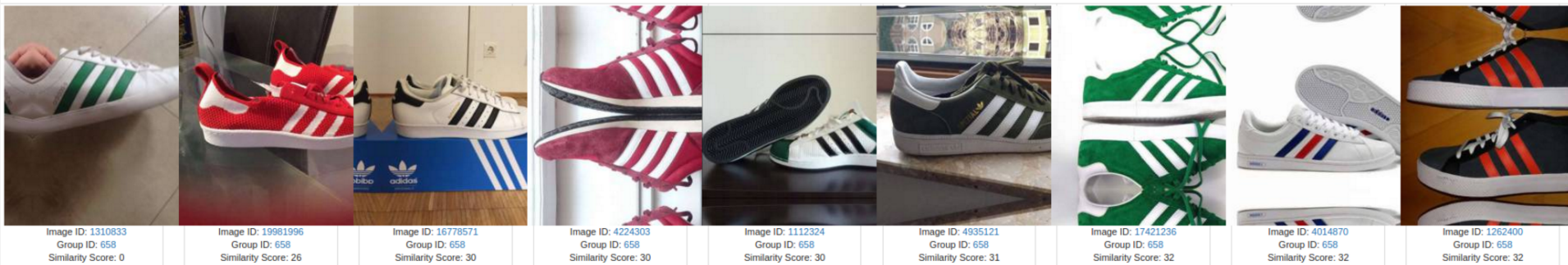
PERFORMANCE QUALITY

- Scalable High-performance cloud solution
- 100% Internal development
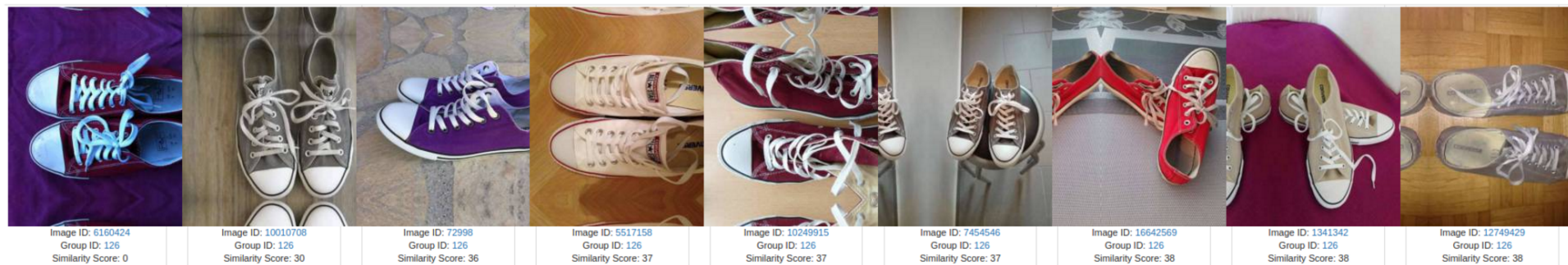- Patent pending: Determining the similarity of images

IN DEVELOPMENT

- Image classification on 800 categories
- Understanding pictures & text
- Brands recognition
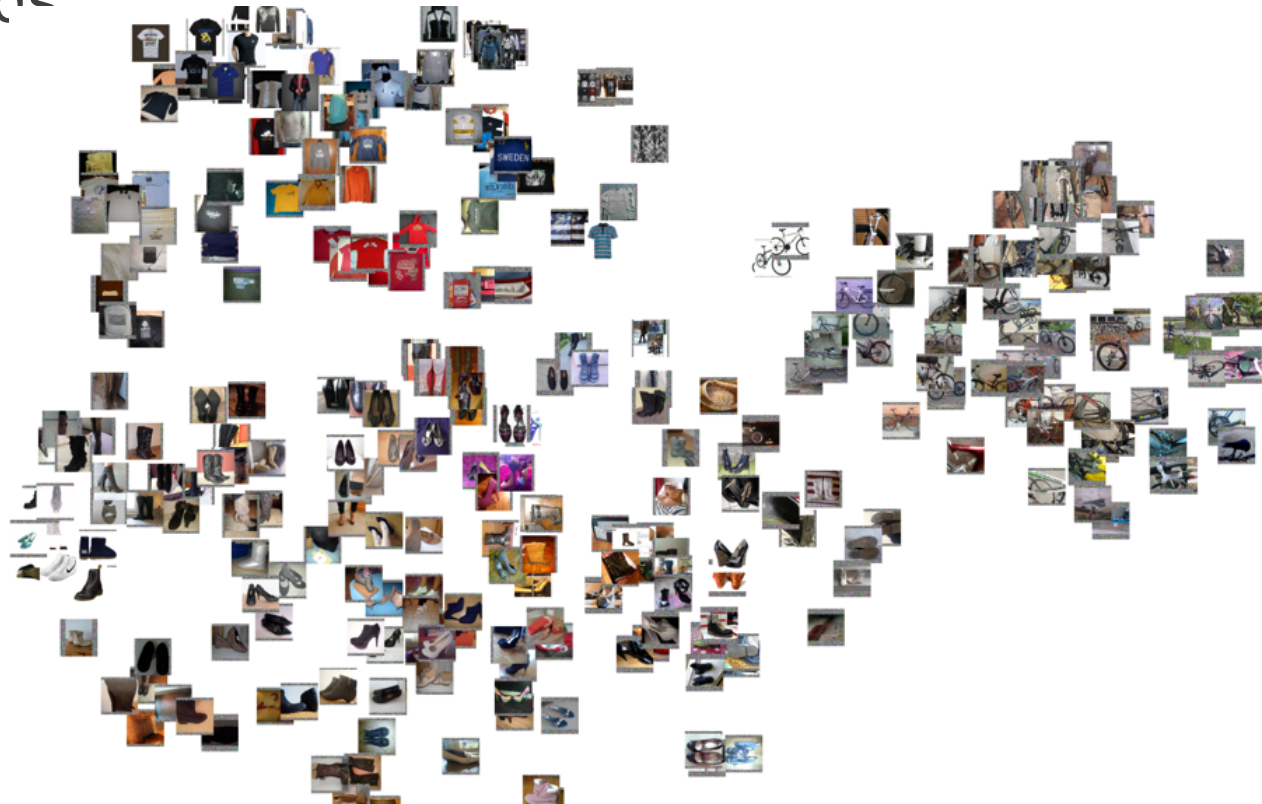
FUTURE PLANS

- New categories & new classifieds
- Categories hierarchy optimisation
- News/Classifieds crosslinking
- Miscategorised ads
- Fraudulent ads
- Duplicate ads
- Customer data

HOP-ON

Why?

- As radical as print to web migration
- Reinforced by the web to mobile migration
- Driven by customer expectations

→ **No classified can miss this train and stay profitable in 5 years!**

Open Discussion !

THANK YOU

Davor Anicic

Business Development and Sales Manager

Styria Digital Services

Styria Medijski Servisi d.o.o.

davor.anicic@styria.hr

ONE SPIRIT,
UNLIMITED IDEAS

(STYRIA