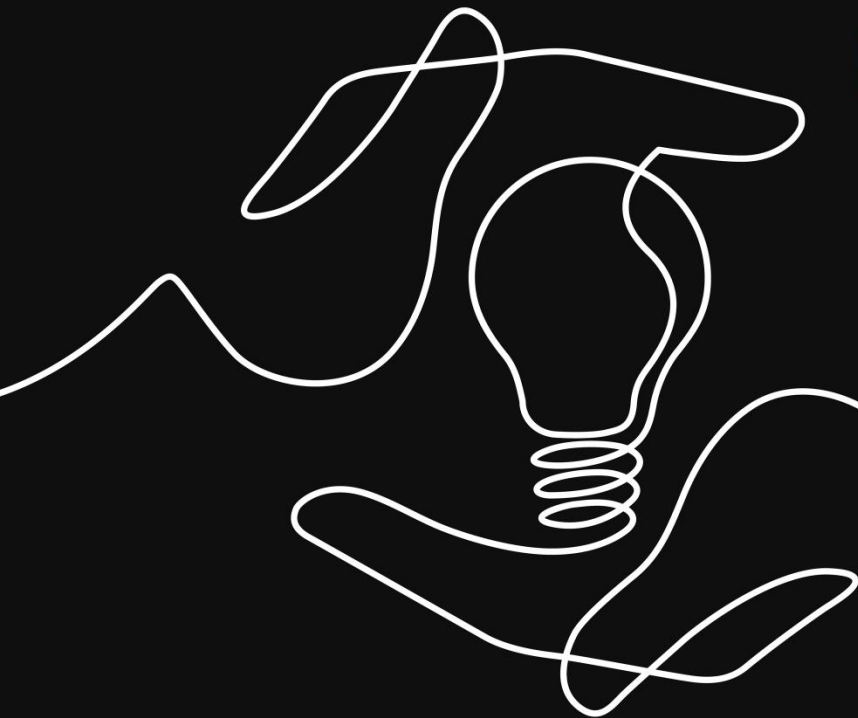


Text Summarization with Transformers

Innovation Week



Innovation
never stops



Atos

Who is?



Florian Luft

- ▶ Senior Expert Data Intelligence
- ▶ Member of global Atos Expert Community
- ▶ More than a decade of Data Analytics experience
- ▶ Focused on
 - semi- and unstructured data
 - Natural Language Processing
- ▶ Public Sector (Defense, Intelligence & Law Enforcement)

Kurz und bündig



107.000 Business Technologists
in 71 Ländern weltweit



#2 in Managed Security
Services weltweit



€11 Mrd. Umsatz und
€1 Mrd. Operative
Marge



ca. €235 Mio. Invest
F&E
jährlich



Weltweiter IT-Partner
der Olympischen und
Paralympischen Spiele



85.000 neue
digitale
Zertifizierungen



Global Leader in **C**loud
und **D**igital Workplace



14.9 tCO2/Mio.€ Umsatz
und damit Branchenführer

Atos Deutschland auf einen Blick

Zwei Headquarters – Bezons (Frankreich) und München

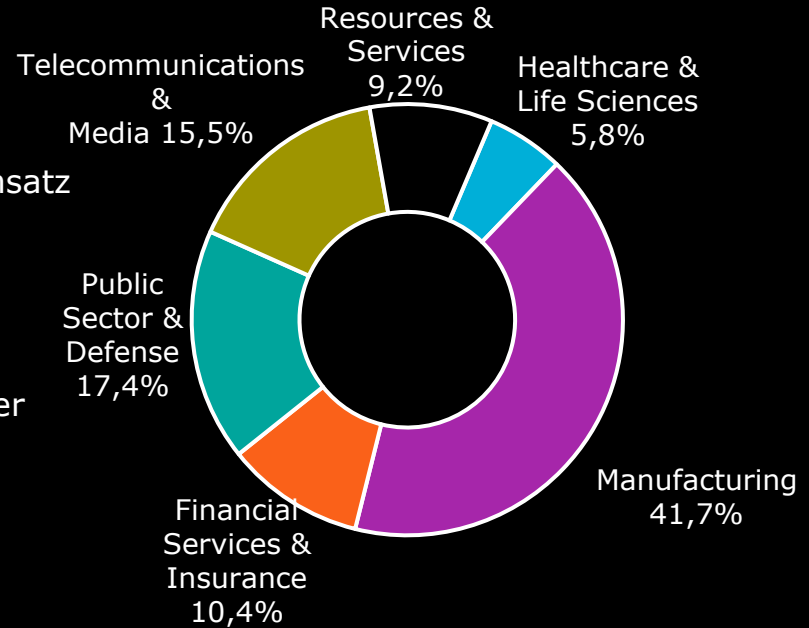


Ca. 9.000 Beschäftigte
(107.000 weltweit)

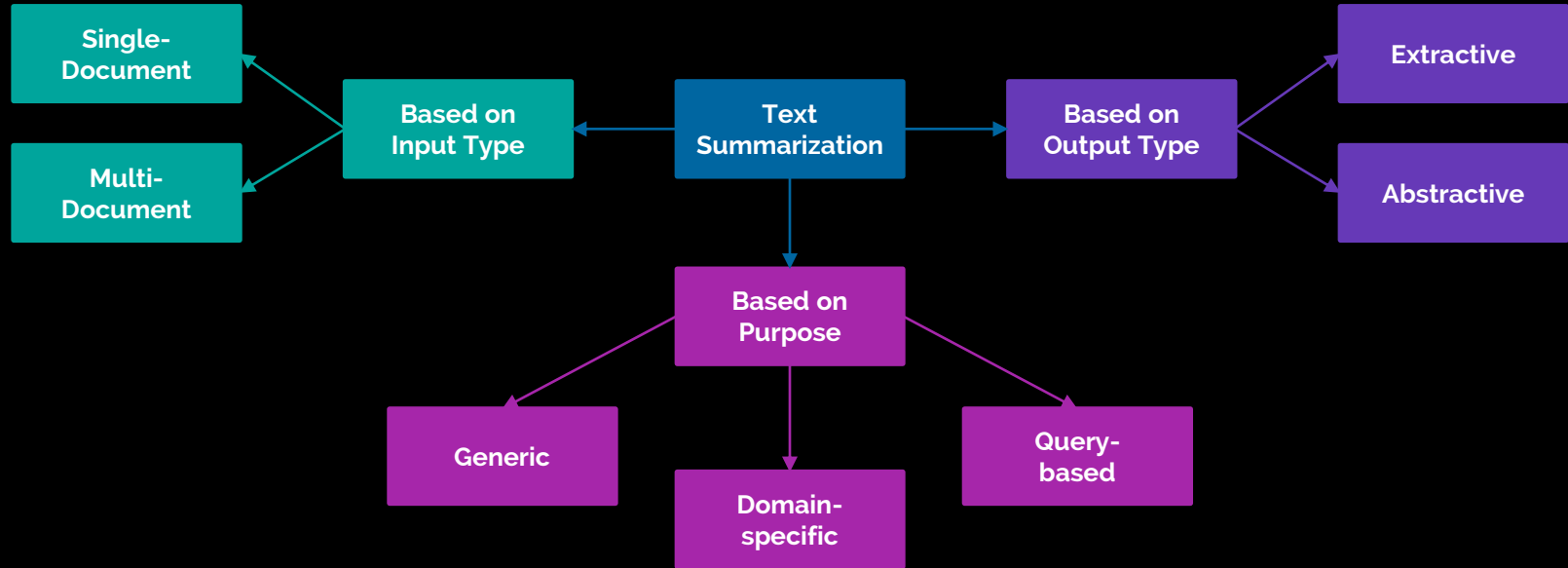
2020: ca. € 2,7 Mrd. Jahresumsatz
in der Region Central Europe
(€ 11 Mrd. weltweit)

München= Globales Headquarter
neben Bezons/Paris

Weitere Standorte:
Berlin, Essen, Frankfurt, Fürth,
Hamburg, Paderborn ...



Text Summarization – Short Intro



Facing the Reality and how Transformers can help

Challenges

- ▶ Abstract Generation
- ▶ Language Dependencies
- ▶ Domain-specific Language
- ▶ ...

Goals

- ▶ Similar to Human-created Summaries
- ▶ Less Training efforts for Models
- ▶ Less Training data requirements

Advantages of Transformers

- ▶ Ability to generate Text rather than extract it
- ▶ Pre-trained for various Languages and Domains
- ▶ Generation of Summaries close to Human-created ones

Transformer Models

BERT (= Bidirectional Encoder Representation from Transformers)

- ▶ Invented by Google
- ▶ Works good for Text Understanding tasks
- ▶ Lacks for Text Generation tasks

GPT-2/3 (= Generative Pre-Training)

- ▶ Invented by OpenAI
- ▶ GPT-3 exclusively licensed to Microsoft
- ▶ Works good for Text Generation

BART (= Bidirectional and Auto-Regressive Transformers)

- ▶ Invented by Facebook
- ▶ BERT-like encoder and GPT-like decoder
- ▶ Best of both worlds: BERT's Text Understanding capabilities + GPT's Text Generation capabilities

T5 (= Text-to-Text Transfer Transformer)

- ▶ Invented by Google researchers
- ▶ Open Source maintained by Huggingface
- ▶ BERT-like encoder and GPT-like decoder

BART and T5 are models of choice for Text Summarization

Use Cases

Market Intelligence

- ▶ Summarization of key competitor activities
- ▶ News tracking
- ▶ Patent research

Financial Research

- ▶ Summarizaion of financial documents like earning reports

Media

- ▶ Generation of newsletters based of full articles
- ▶ Teaser generation
- ▶ Generation of condensed content for mobile device use

Intelligence / Defense

- ▶ Summarization of analyst reports and open sources

Pharma

- ▶ Summarization of research papers and trials during clinical research phase

Thank you

For more information please contact:

T+ 49 211 39912862

M+ 49 160 4904733

florian.luft@atos.net

Atos, the Atos logo, and Atos|Syntel are registered trademarks of the Atos group.
April 2021. © 2021 Atos. Confidential information owned by Atos, to be used by the recipient only. This document, or any part of it, may not be reproduced, copied, circulated and/or distributed nor quoted without prior written approval from Atos.

The Atos logo, featuring the word "Atos" in a stylized blue font. The letter 'o' is replaced by a circular graphic element.