Have your cake and eat it

How the Magenta Digital Assistant reconciles data-driven AI with privacy protection

Harald Störrle
2022-11-03

Abstract

Traditionally, data privacy is considered a run-time topic: how to keep operational data safe and use it only in admissible ways. Testing, in contrast, happens at build-time, and uses test data, thus all but excluding privacy concerns altogether. For AI-based systems, the distinction between operational data and test data all but disappears.

The underlying reason is, of course, that deep learning algorithms hinge on the availability of large amounts of labelled data - and the only meaningful data is *your* data. This seems to be a tough choice between using AI and keeping data privacy. Can't we have both?

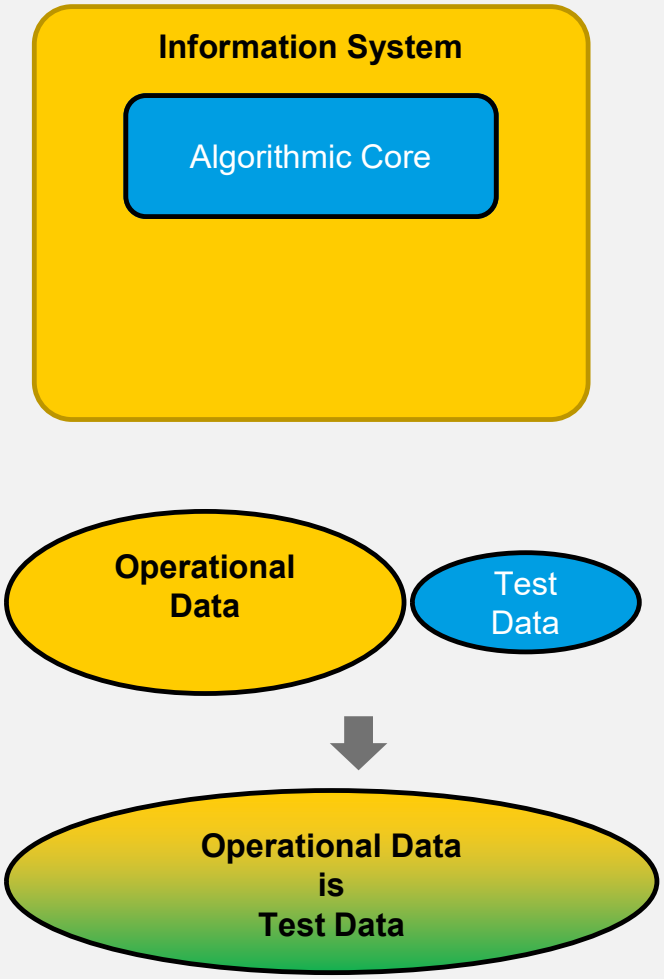
Yes, you can - as we have demonstrated with "Hallo Magenta", a state-of-the art digital assistant by Deutsche Telekom. It powers the SmartSpeakers, Smartphone Assistant Apps, and Set Top Box Remote Control Units by Deutsche Telekom and Orange France.

In this talk, I will explore the consequences of building and operating AI-based systems with regards to GDPR-compliance. I will first briefly outline data driven AI, the General Data Protection Regulation (GDPR), and how they may clash. Then, I will show how AI and GDPR can be reconciled using the digital assistant "Hallo Magenta" by Deutsche Telekom as a case study. Time allowing, I will conclude with an outlook into why digital assistants are the key to the digitalization of public services in Germany.

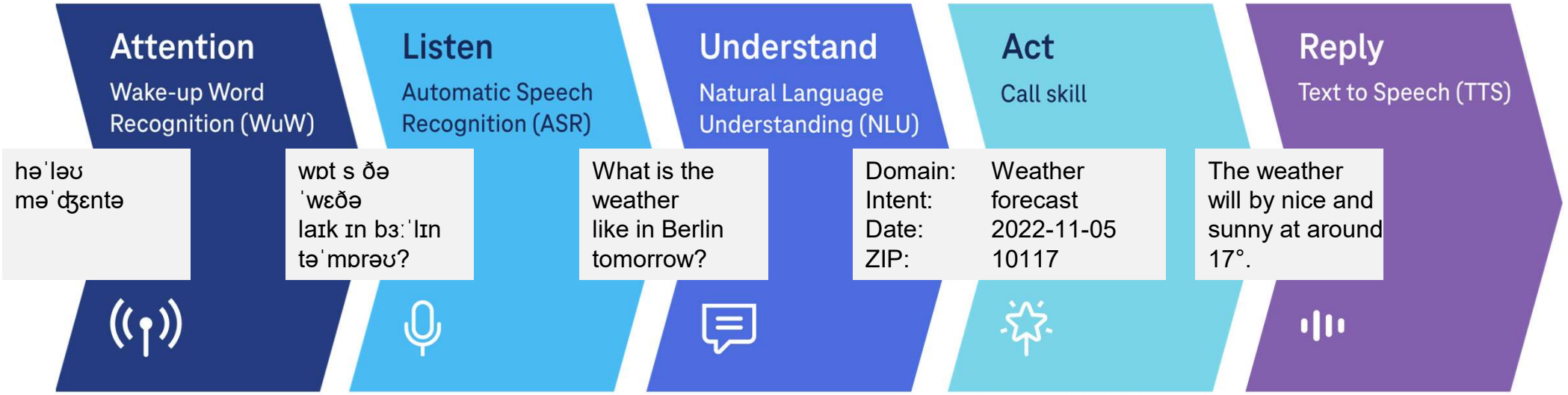
What is an "AI System"?

- An "AI System" is really a plain old Information System, as we have done since the 1950s, …
- … that replaces and/or complements ist algorithmic core by one or more "AI" cores to achieve their functional goal.
- "AI" really just means using data-driven statistical machine learning, often of the deep learning variety.

- Observations
  - All established challenges of Information Systems construction remain.
  - Genuine ML challenges on top, e.g. data acquisition, data quality, training, overfitting, …
  - The distinction between operational data at run-time and test data at build-time all but collapses -- safeguarding privacy on test data is the new main challenge.



Hallo Magenta on TVS
Overall Functionality



Activate listening by trigger, such as wake-up phrase or button push

Turn speech input to written text

Identify intents and entities in text

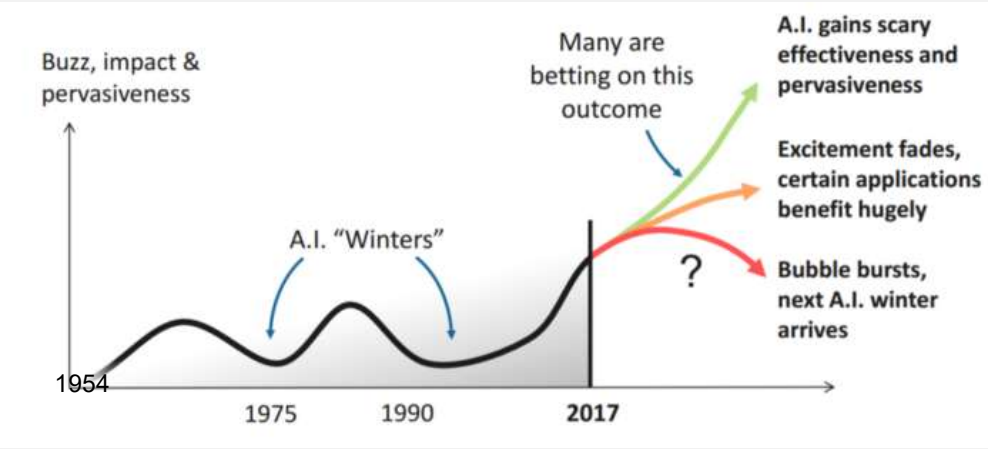
Act on the understood content by adjusting settings of speaker (e.g., volume), accessing a data base (e.g. Wikipedia), or controlling a SmartHome device.

Convert the textual skill result into spoken words (or render content on a screen)

Artificial Intelligence

- 1954: "Artificial Intelligence" coined as a marketing catchphrase at Dartmouth summer school
  - symbolic vs. sub-symbolic approaches compete
  - After limits to ANN become obvious with the Perceptron, symbolic approachs mostly prevail
  - sequence of boom-and-bust-cycles ("AI winter")
- 2000s: Deep learning breakthrough due to three factors
  - massive computing power (aka. GPUs, TensorFlow)
  - availability of massive labelled data datasets
  - faster convergence in training
- Turing Award 2018 to Bengio, LeCun, Hinton, the "Fathers of the Deep Learning Revolution"

https://awards.acm.org/about/2018-turing
Yoshua Bengio, Yann Lecun, Geoffrey Hinton: Deep learning for AI. Com. ACM, 64 (7) July 2021 pp 58–65, https://doi.org/10.1145/3448250



https://medium.com/predict/what-happens-when-ai-is-let-out-of-our-boxes-8505e17ba00d
https://www.techtarget.com/searchenterpriseai/definition/AI-winter

Data Privacy

- Privacy was elevated to a basic constitutional right in Germany with the 1983 "Volkszählungsurteil".
  - Today, privacy is considered a universal human right on a par with freedom of speech (UN Universal Declaration of Human Rights, §12).
- German law was (more or less) generalized to the EU by the 2018 General Data Protection Regulation (GDPR, aka. DS-GVO).
  - Applies to all people in Europe - no matter their nationality, the affiliation of processor, or country of processing.
  - Subsequent regulations in California, Japan, Korea very similar (cf. CCPA, PIPA, APPI).
  - The GDPR text is readily available: several languages (EN, DE), surprisingly easy to read.
- After grace period, regulators now increasingly enforce rules, see enforcementtracker.com.
  - Fines of up to 2% annual global turnover.
  - Fines of up to 4% a.g.t. when failing to implement regulators guidelines, or repeat offenses.
  - Loss of brand value/customer trust may be even larger - trust is hard won and easily lost.
- In Germany, there is generally a high level of interest and sensitivity in this topic.
  - Several active parties (e.g. CCC, heise, Handelsblatt, golem, Ges. für Freiheitsrechte)

GDPR Article 5
Principles relating to processing of personal data

Personal data shall be:

- processed lawfully, fairly and in a transparent manner in relation to the data subject **('lawfulness, fairness and transparency')**;
- collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes **('purpose limitation');**
- adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed **('data minimisation');**
- accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay **('accuracy');**
- kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject **('storage limitation');**
- processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures **('integrity and confidentiality')**.

The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 **('accountability').**

Privacy and Smart Speakers in the press

4. Oktober 2019, Moritz Tremmel, Golem.de
Das allgegenwärtige Ohr Amazons
Die kürzlich angekündigten Echo-Produkte bringen Amazons Sprachassistentin Alexa auf die Straße und damit Datenschutzprobleme in die U-Bahn oder in bisher Alexa-freie Wohnzimmer. Mehrere Landesdatenschutzbeauftragte haben Golem.de erklärt, ob und wie die Geräte eingesetzt werden dürfen.

4. Juli 2019, Oliver Nickel, Golem.de
Alexa und Drittanbieter speichern Daten unendlich lang
Erst wenn Kunden aufgenommene Daten manuell auswählen, werden sie gelöscht. Das bestätigt Amazon auf eine Anfrage des US-Senators John Coon. Selbst dann seien einige Ausnahmen zu beachten, in denen Alexa und Anbieter Daten weiterverarbeiten.

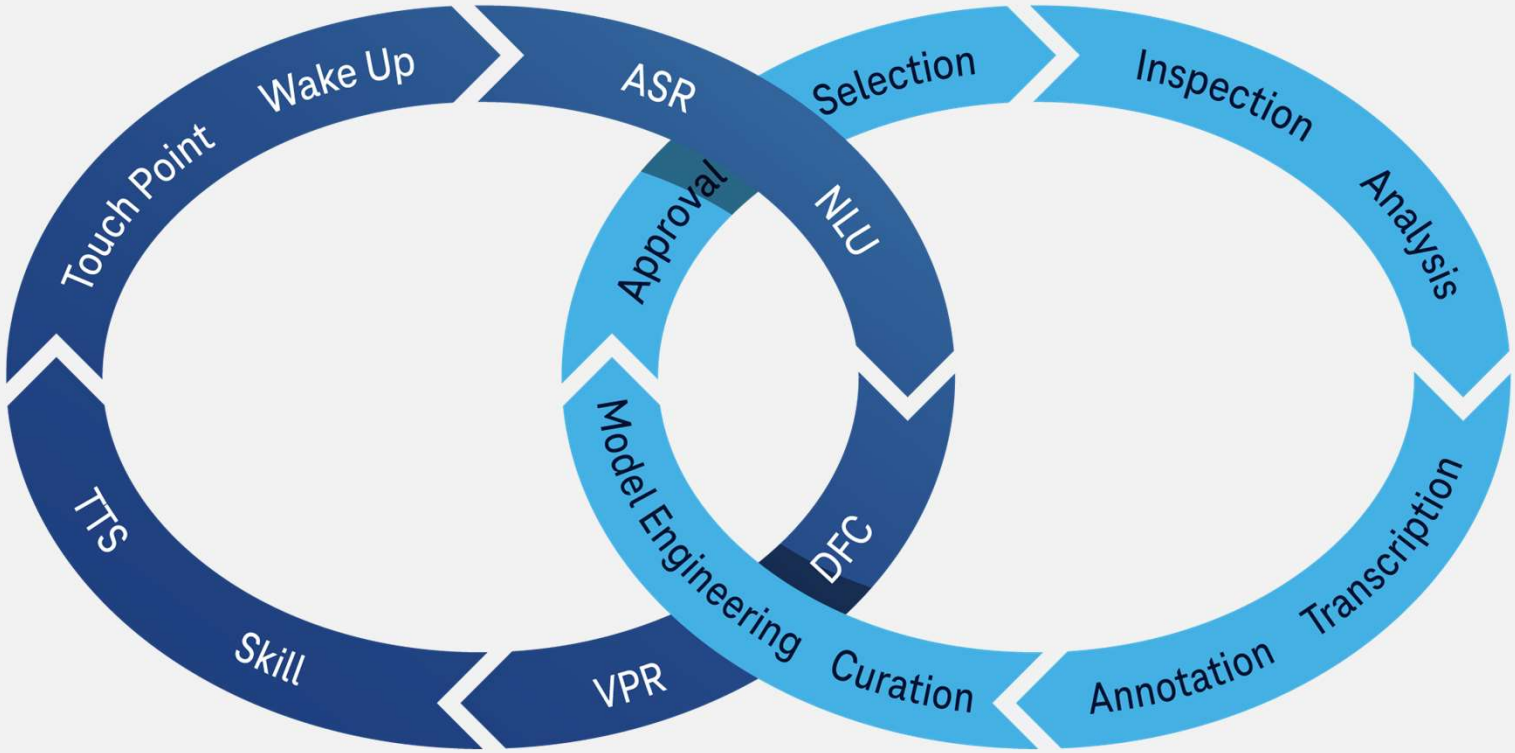
11. April 2019, Ingo Pakalski, Golem.de
Mitarbeiter sehen Alexa-Befehle mit verknüpften Kundendaten
Amazon-Mitarbeiter lauschen aufgenommenen Alexa-Sprachbefehlen - und erhalten dazu das passende Amazon-Konto samt Vorname des Kunden. Damit können Menschen die Sprachbefehle in Grenzen bestimmten Personen zuordnen. Apple und Google versprechen einen besseren Schutz der Privatsphäre.

7. Juli 2019, Oliver Nickel, Golem.de
Apple hört durch Siri Drogengeschäfte und Sex mit
Einige Apple-Vertragsarbeiter erhalten Sprachdateien von Siri, um diese zu analysieren und den Assistenten zu verbessern. Allerdings werden laut einer internen Quelle des Guardian oft private Momente aufgenommen - beim Sex, bei Geschäften und beim Gespräch mit dem Arzt.

26. Juli 2019, Alex Hern, The Guardian
Apple contractors 'regularly hear confidential details' on Siri recordings
Workers hear drug deals, medical details and people having sex, says whistleblower

11. Juli 2019, Moritz Tremmel, Golem.de
Google-Mitarbeiter hören Gespräche mit
Jeder 500. Sprachbefehl wird von Google-Mitarbeitern ausgewertet. Ein Leak aus Belgien zeigt, dass sich darunter auch viele ungewollt aufgenommene Gespräche und Telefonanrufe befinden - mit privaten und intimen Inhalten.

Telekom Voicification Suite (TVS)
More than voice processing



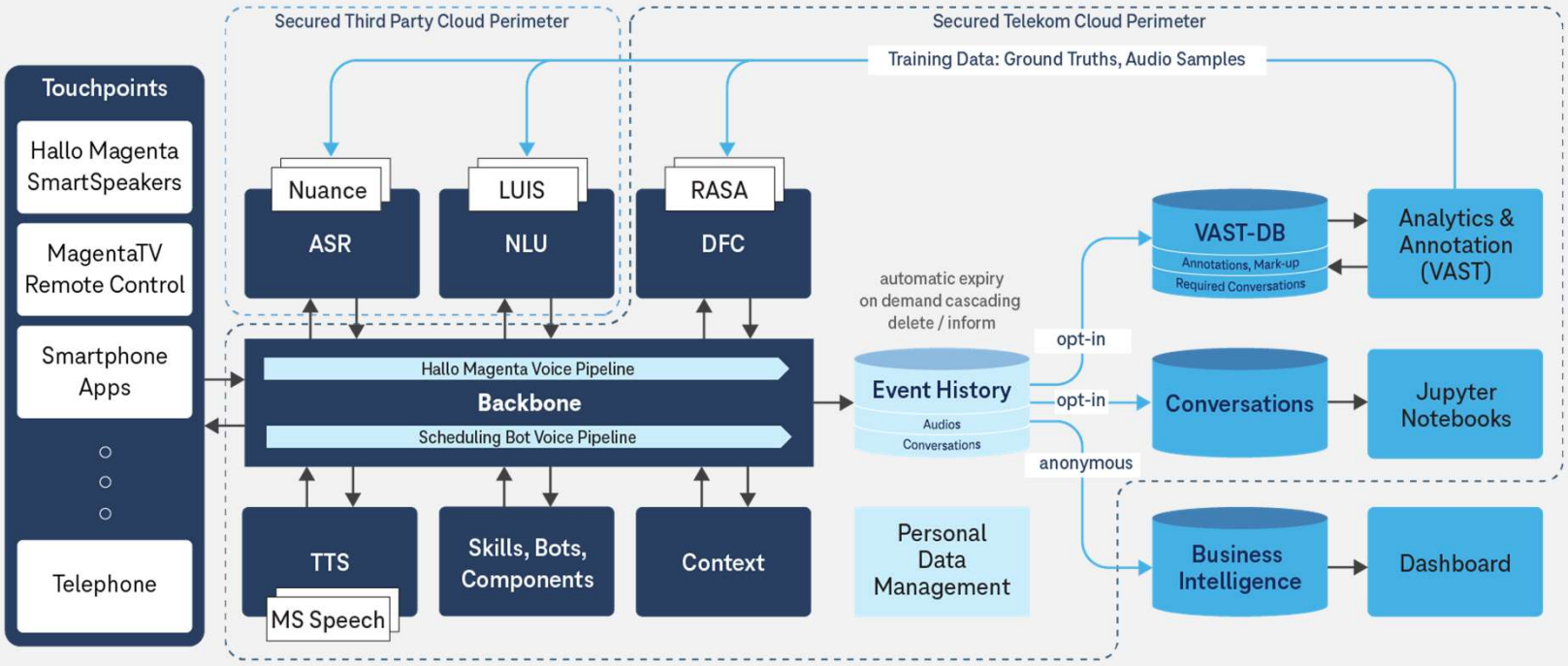
Voice processing contains two parts

- Everybody knows about ASR, NLU, and skills - cf. Alexa, Echo, Cortana, Siri, Bixby, …
- However, The actual voice processing is but the visible part of the platform.

Industrial applications

- For industrial applications, the support processes are equally important - if not more so!
- The reason behind this is simple: the core of a voice based application are the conversational capabilities and the interaction quality, which crucially depend on the AI models, and thus on the data, their quality, and the data curation pipeline.
- Without such a pipeline, and the appropriate tooling, voice bsed applications remain prototypes, not industrial strength applications.

Telekom Voicification Suite (TVS)
More than a system



Industrial strength platform

- Multi-tenant capable
- High quality code
- Low(er) CO2 footprint

Flexibility built into the platform

- Multiple alternatives for WUW, ASR, NLU, Dialog flow-control, TTS
- Customizable voice processing pipelines
- optional components for VID, Gender rec., DCS, MALA, NEMS, SLRR, IVR, …

Business-ready solution

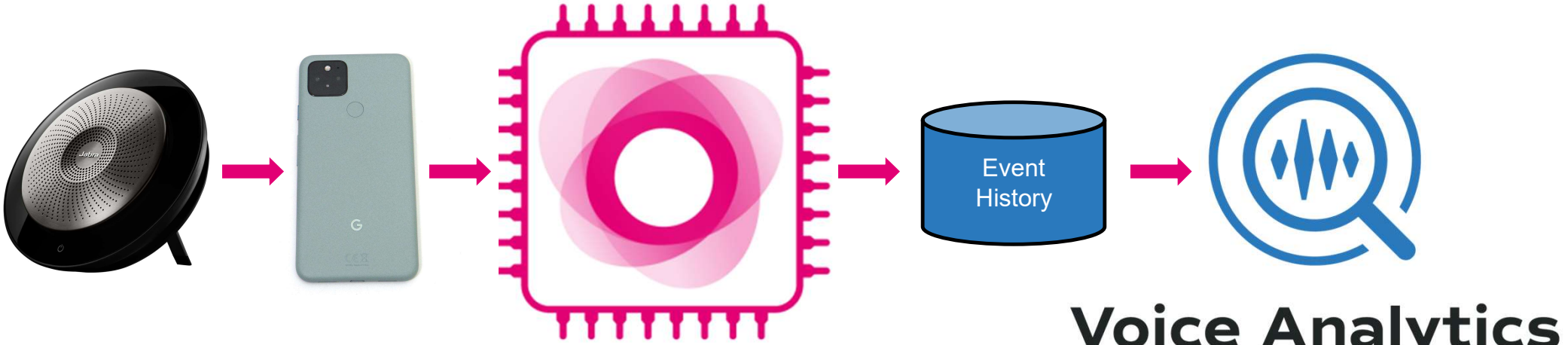
- Support for AI training & optimization
- Support for Monitoring & debugging
- Support for Analytics, BI, Data Science
- Support for business processes (PDM, Billing, …)

GDPR compliance from the start

What does the TVS do to ensure Privacy?

- "Straightforward" technical measures
  - Servers in EU
  - encryption
  - 2FA
- Legal and organisational measures
  - Legal constraints on sub-processors & annotators
  - Annotators from TK only, on-site, TK devices only
  - Annotator enablement, supervision & qualification
  - Restrictive rights, on/off-boarding processes
  - Internal assessment process (PSA)
- Privacy by design
  - Opt-in only
  - Data expiry (automatic + on-demand)
  - All data labelled by sensitivity
  - Transparency (e.g. §15)
  - Access protocol (encrypted)
  - Masking
  - Closed circuit training

Terminfinder Demo



☎ +49 ### #### ##13

Core Benefits

- Cost reduction (-42%)
- increased employee productivity (-22.5% time)
- 24/7 availability (+82.5%)
- no wait time for customers
- no lost calls due to busy lines
- GDPR compliant

Public demonstrations by leading Telekom personnel

- https://www.linkedin.com/posts/magenta-voice_the-end-of-waiting-time-in-call-center-hotlines-activity-6945609262115004416-43_v?utm_source=linkedin_share&utm_medium=member_desktop_web
- https://www.linkedin.com/feed/update/urn:li:activity:6942772122570121216/
- https://www.linkedin.com/feed/update/urn:li:activity:6990922393267400705
- https://geschaeftskunden.telekom.de/magenta-business-collaboration/ki-basierte-bots
- https://geschaeftskunden.telekom.de/magenta-business-collaboration/conversational-ai/terminfinder

QA|WARE

SOFTWARE ENGINEERING

QAware GmbH

Aschauer Straße 32

81549 München

Tel. +49 89 232315-0

info@qaware.de