

Virtual Therapy in Head&Neck cancer

Data collection and sharing for articulatory synthesis of patient speech

Rob van Son

Netherlands Cancer Institute, Amsterdam, The Netherlands
R.v.Son@nki.nl

15 November 2017



Introduction



TAPAS: Training Network on Automatic Processing of PAtiological Speech

EU promotes Open Data and Open Science

- TAPAS will collect unique data on speakers and pathological speech
- Speakers are “vulnerable”: patients and children
- Most stringent requirements on privacy protection (GDPR&CTR)
- But we want to share

Collections of Patient Speech and PROs



Current practices regarding speech

Data of Head&Neck tumor patients

Protocols: Speech and Patient Reported Outcomes (PROs¹)

- Data pre- and post-treatment (6mnd+12mnd)
- Oral Cavity protocol (~95)
 - Sustained a:u:-i:; ei-au-œy; word-list (36w); story[†] (75w); pa-ta-ka
 - PROs: SHI, SWAL QoL, EORTC QLQ-H&N35
- Larynx protocol (~150)
 - a: Longest/high/low/loud/soft/sweep; story[†]
 - PROs: LASA, EAT-10, VHI, EORTC QLQ-C30/H&N35
- Tracheolaryngectomy (TLE) protocol (~25)
 - a: Longest/high/low/loud/soft/sweep; story[†]; 3 voiced sentences*
 - PROs: EQ-5D-5L, SOAL, VHI-10,

¹PROs are questionnaires asking about health and QoL

Secondary use of patient data

Currently *opt-out*

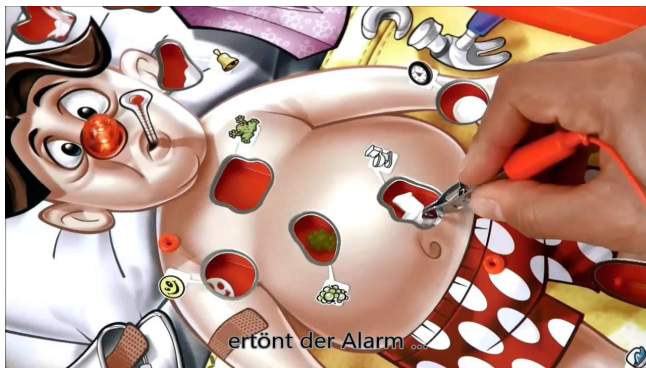
- Health care data, will switch to *opt-in* (informed consent)
- Restricted use (*severely* restricted under opt-out)
- IRB approval needed for *each* project
- Can we create a speech “Biobank” from secondary use data?
Probably not

Switch to *Informed Consent*

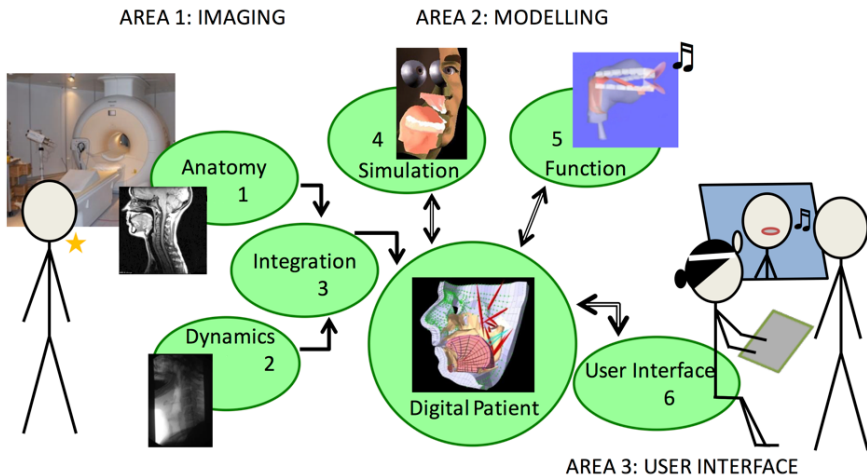
- All patients asked to consent
- Look for broad consent
Unclear whether broad consent is possible

Virtual Therapy

Predicting and synthesizing plausible speech examples after oral cancer treatment



Demonstrating functional outcome of therapy



Data collection

Digital Patient

1 Anatomy & Physiology

MRI, DTI-MRI, Shear-Wave US elasticity
3D Photo of tongue shape and resection

2 Dynamics

dynamic MRI, Tongue mobility&strength,
3D video, sEMG

3 Integration

ArtiSynth bio-mechanical model

4 Simulation

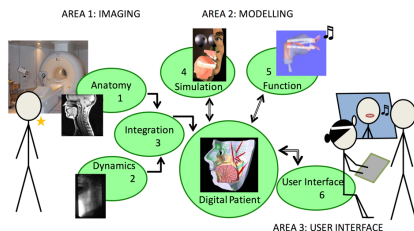
Forward and inverse model training

5 Function (Speech)

Real Speech & Articulatory synthesis

20 patients + 5 healthy

- 10 small tumors in oral tongue (surgery)
- 10 larger tumors in base of tongue (radio-therapy)



Data Sharing



What can be shared, when, and with whom?

Talk to:

MEC, Privacy Officer, Biobank organizers, Privacy lawyers

Confusion about *Informed Consent* (IC) in Open Data/Science

- IC determines what can be done with the data, i.e., cover all uses
- Confusion: IC covered by GDPR and CTR² [1, 2]
 - GDPR: Consent must be specific [3, 4]
 - CTL: One-stop informed consent possible [5, 6]
- EU vs. National rules on health data and consent (CTR, [4])
- What health data fall under the research derogation of GDPR, if any?
- What research is “in the public interest”?
- Open data is international, the GDPR restricts cross-border exchange

²Clinical Trials Regulation (536/2014)

Consensus: Privacy by design

Demands on shared data (under the GDPR)

- Data minimization *what is not there, cannot be exposed*
 - Coarse-graining: age-brackets, truncate zip codes, etc.
 - Strip metadata from images, movies, MRI
 - Censor bars in pictures, movies, MRI
 - Anonymization *if data is useful, it is not anonymous*
 - Pseudonymization
 - Encryption
 - Security, computer and otherwise
 - Data transfer agreements, NDA's, Promise of Confidentiality
- ⇒ Take the analysis to the data *privacy-protecting platform*

Take the analysis to the data

On-demand analysis from a web browser

- Only a single DTA contract between data-owner and platform owner
- Analyse data without access to micro-data
- Users only see the outcome of the certified analysis
- Platform supplies tools: e.g., [R](#), [Bob and Kaldi](#) [7, 8]
- Platform enforces access rights, audit trails, and security

The BEAT platform

[10]

Biometrics Evaluation and Testing (BEAT)

[9]

- EU 7th framework program
- Part of European computing e-infrastructure for Open Science
- Solution for open access, scientific information sharing and re-use
- Data and source code
- Protecting privacy and confidentiality
- Data from different experiments can be easily compared and searched
- Challenges and education
- Attestation mechanism for reports
- Chosen as the data platform for TAPAS

Anjos et al. (2017). *BEAT: An Open-Source Web-Based Open-Science Platform*. arXiv preprint arXiv:1704.02319.

BEAT overview

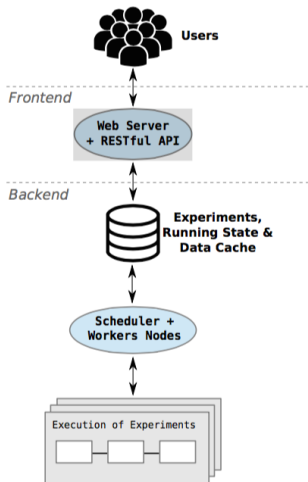
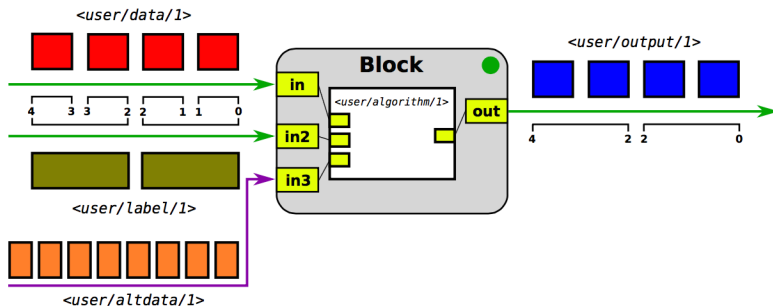


Figure 1: An overview of the BEAT platform applications and their interaction. Users use the web front-end to run experiments, search and combine results. A back-end handles the execution of experiments on dedicated hardware.

BEAT toolchain block



Individual Blocks are strung together into tool-chains

BEAT experiment configurator

Name:
 ✓
 Enter a meaningful name to help you recognize this experiment. Auto-completion will help you in keeping your naming conventions tidy. If a chosen name is **highlighted in red**, it is because it is already being used. In this case, choose another name.

Toolchain:
[tutorial/eigenface/1](#)

Datasets [Advanced mode]

Database

Protocol

Blocks

linear_machine_training	<input type="text" value="Select an algorithm..."/>
template_builder	<input type="text" value="Select an algorithm..."/>
probe_builder	<input type="text" value="Select an algorithm..."/>
scoring	<input type="text" value="Select an algorithm..."/>

Analyzers

analysis	<input type="text" value="Select an algorithm..."/>
-----------------	---

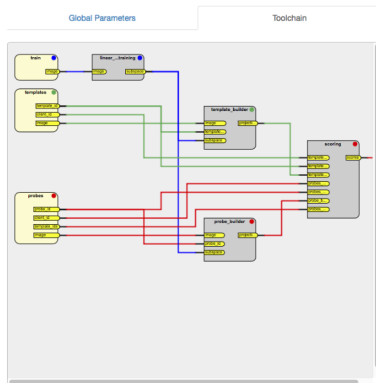
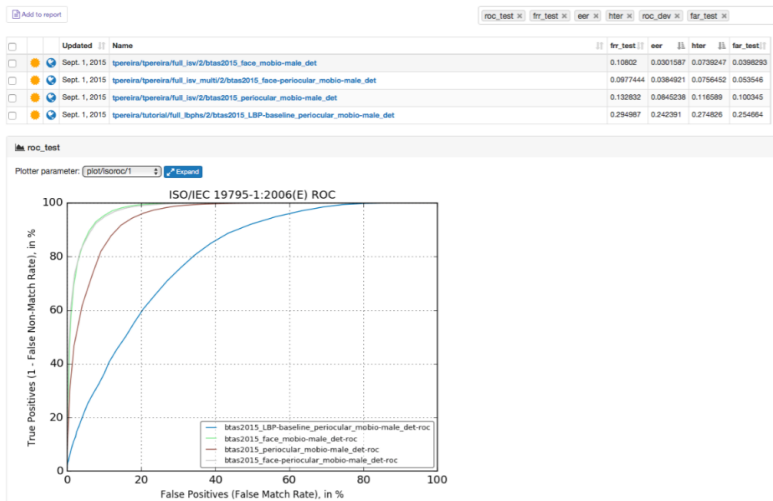


Figure 3: The BEAT platform experiment configurator allows the user to easily associate databases, algorithms and analyzers together to create the desired setup. As the user chooses components for the toolchain, choices of further components are restricted respecting data format compatibility between the blocks.

BEAT automatically generated tables



Thank You!

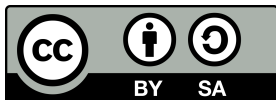
?

More information I

- [1] European Parliament and the Council of the EU, "Clinical trials regulation 536/2014." https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2014_536/reg_2014_536_en.pdf, 2014.
- [2] C. K. Schneider *et al.*, "Deciphering the eu clinical trials regulation," *Nature biotechnology*, vol. 34, no. 3, p. 231, 2016.
- [3] European Union, "General Data Protection Regulation, GDPR." <https://gdpr-info.eu/>, 2016.
- [4] G. Chassang, "The impact of the eu general data protection regulation on scientific research," *ecancermedicalsecience*, vol. 11, pp. 709, doi:10.3332/ecancer.2017.709, 2017.
- [5] A. Mende, M. Frech, and C. Riedel, "Principles of the eu clinical trials regulation no 536/2014: What will change?," *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 2017.
- [6] C. Dittrich, A. Negrouk, and P. G. a. Casali, "An ESMO-EORTC position paper on the EU clinical trials regulation and EMA's transparency policy: making european research more competitive again," *Annals of Oncology*, vol. 26, no. 5, pp. 829–832, doi:10.1093/annonc/mdv154, 2015.

More information II

- [7] M. Cernak, A. Komaty, A. Mohammadi, A. Anjos, and S. Marcel, “Bob speaks kaldi,” in *Proc. of Interspeech*, no. EPFL-CONF-229211, 2017.
- [8] A. Gaye, Y. Marcon, J. Isaeva, P. LaFlamme, A. Turner, ..., and P. R. Burton, “DataSHIELD: taking the analysis to the data, not the data to the analysis,” *International Journal of Epidemiology*, vol. 43, no. 6, pp. 1929–1944, doi:10.1093/ije/dyu188, 2014.
- [9] Idiap Research Institute, “Biometrics Evaluation and Testing (BEAT) .” <https://www.beat-eu.org/>, 2017.
- [10] A. Anjos, L. El-Shafey, and S. Marcel, “Beat: An open-source web-based open-science platform,” *arXiv preprint arXiv:1704.02319*, 2017.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

©2017 R.J.J.H. van Son