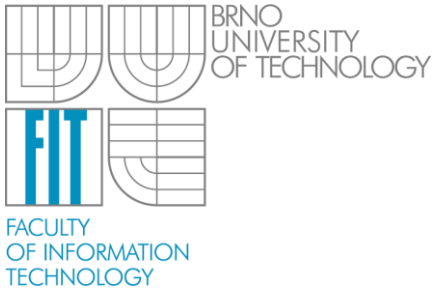
Voice Biometry standard proposal

Honza Černocký

Brno University of Technology,
BUT Speech@FIT,
Czech Republic



Sep 8th 2015, Interspeech VBS meeting

Program

Honza Cernocky – intro, “why?”

Ondrej Glembek – Technical description

Petr Schwarz – Phonexia remarks

Discussion

Honza Cernocky – next steps

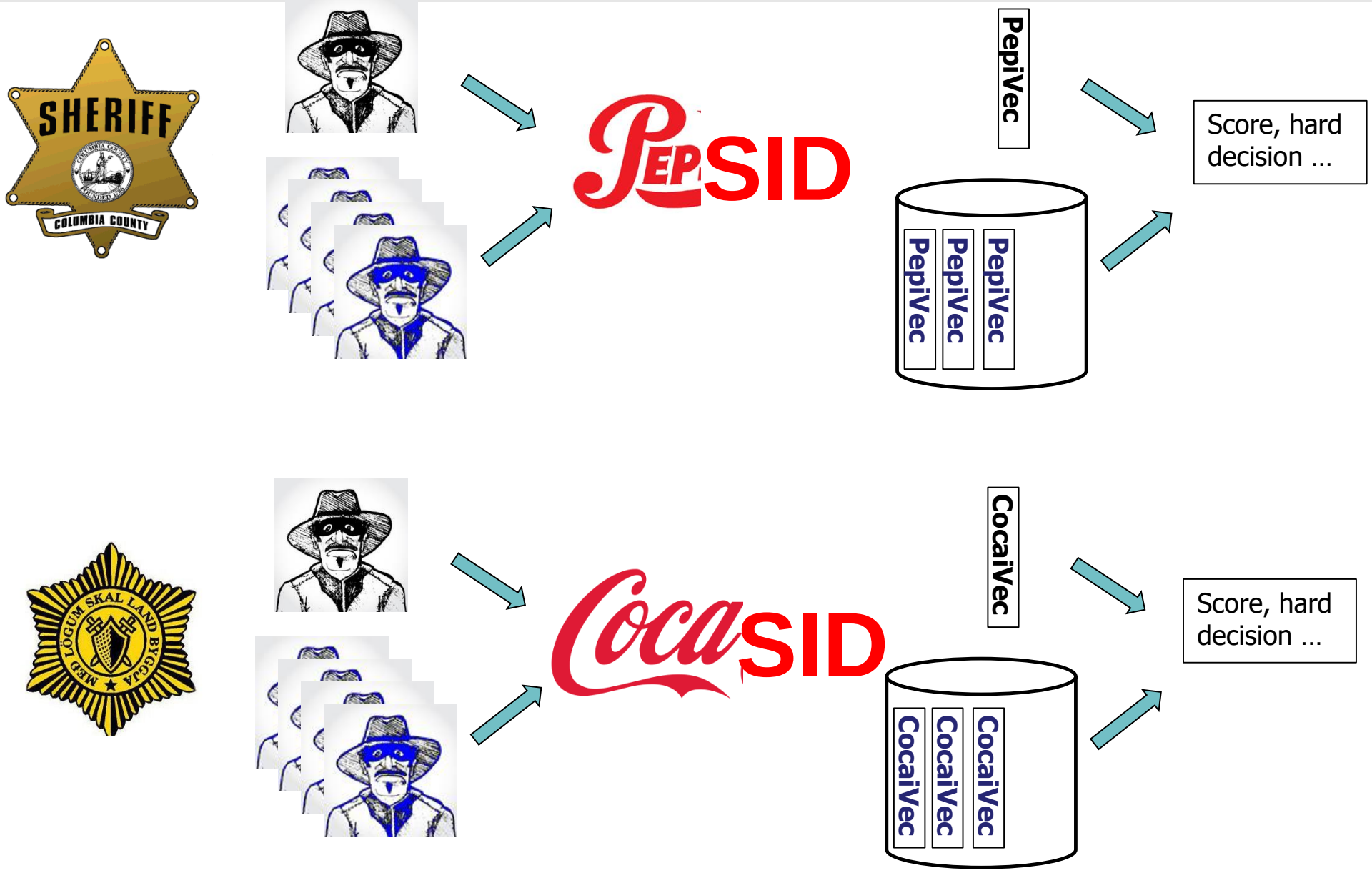
End 16.00, no buffet, drinks, entertainment ☹

Situation

- In the last 10 years, scientific advances in speaker recognition (JFA, iVectors, PLDA) allowed for producing precise and robust SRE systems
- Quickly adopted by vendors, producing solutions that **are successful on the market**.
- R&D never stopping
  - Everyone continuously improving performance of their system, robustness, calibration, etc
  - New versions of engines released

A vibrant community working in cooperative/competitive mode both for R&D labs and vendors.

It works ☺


SHERIFF
COLUMBIA COUNTY
PepSID
PepiVec
PepiVec
PepiVec
PepiVec
Score, hard decision ...
MEÐ LÖGUM SKAL LAND BYGGJA
CocaSID
CocaiVec
CocaiVec
CocaiVec
CocaiVec
Score, hard decision ...

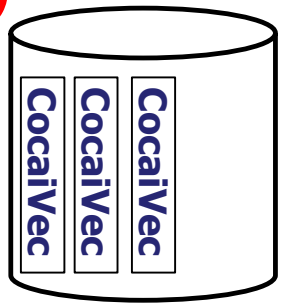
It does not work ☹

PepSID

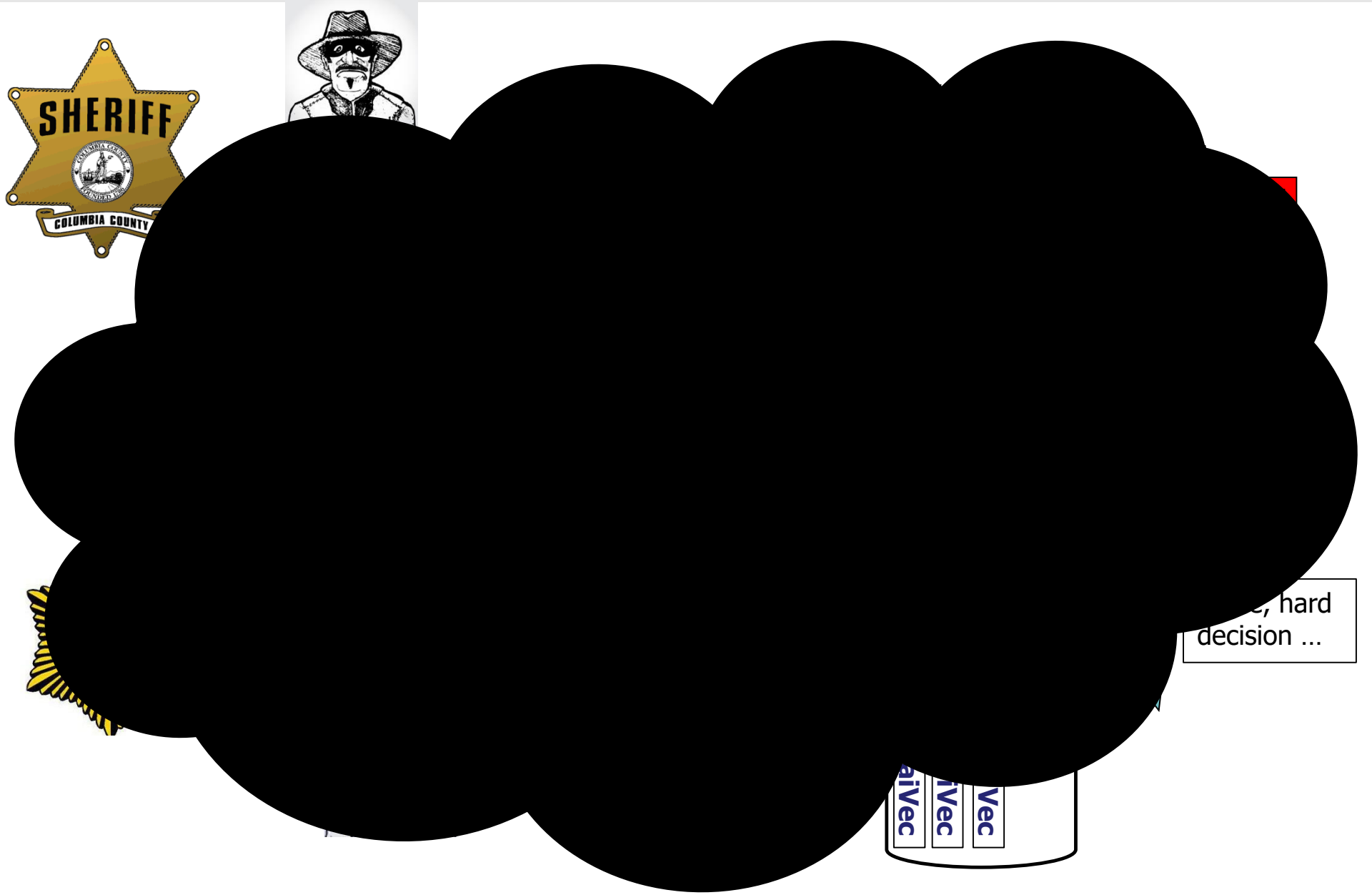
PepiVec

MISMATCH

CocaSID



Making it work



Making it really work – standardized iVectors

PepSID

VBSiVec

Score, hard decision …

CocaSID

VBSiVec VBSiVec VBSiVec

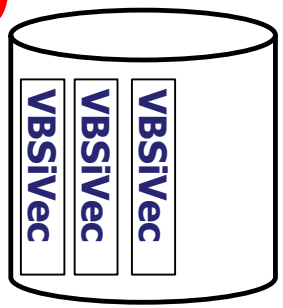
Making it really work – standardized iVectors

PepSID

VBSiVec

Score, hard decision …

CocaSID



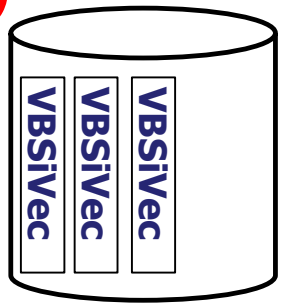
Making it really work – standardized iVectors

PepSID

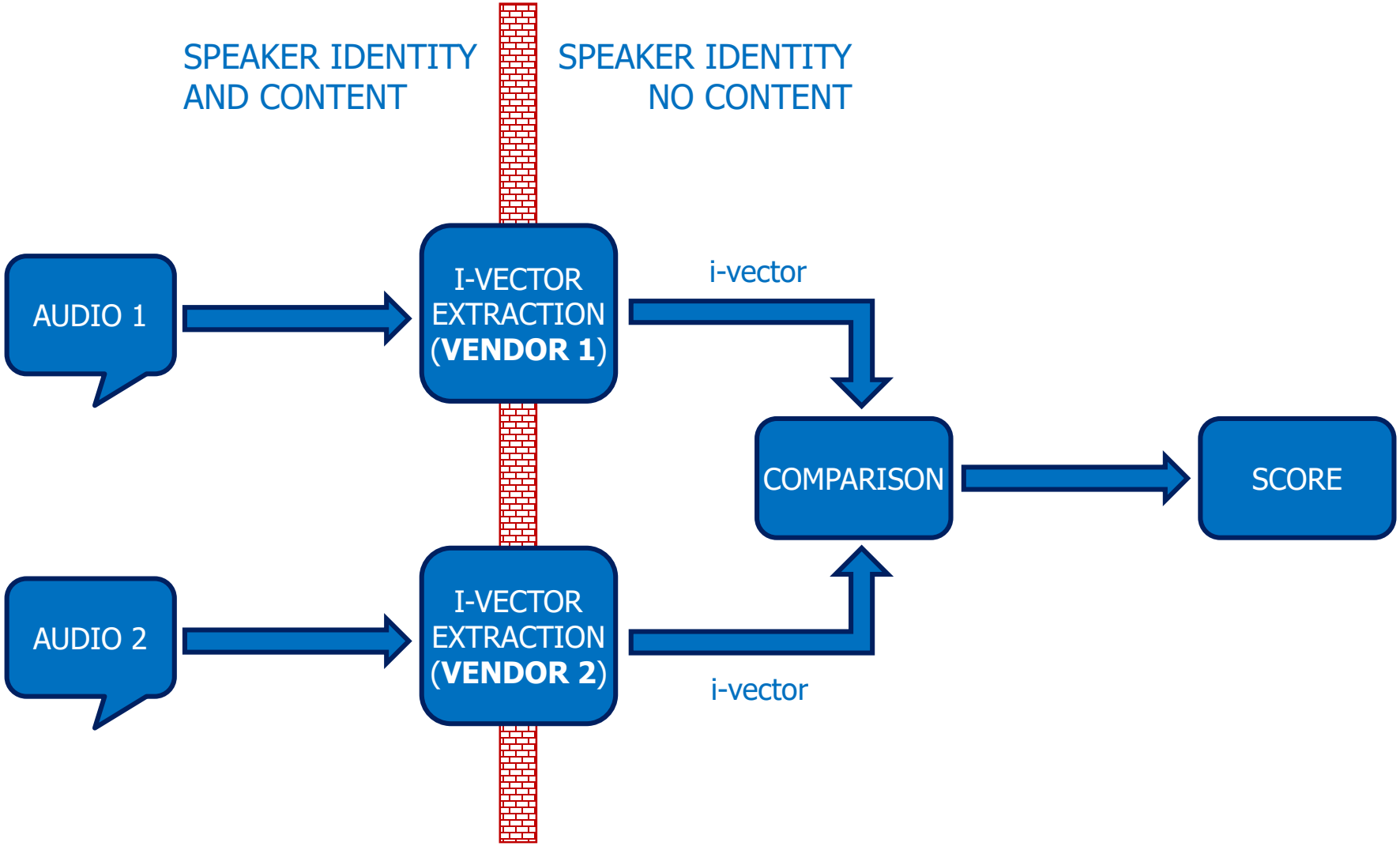
VBSiVec

Score, hard decision …

CocaSID

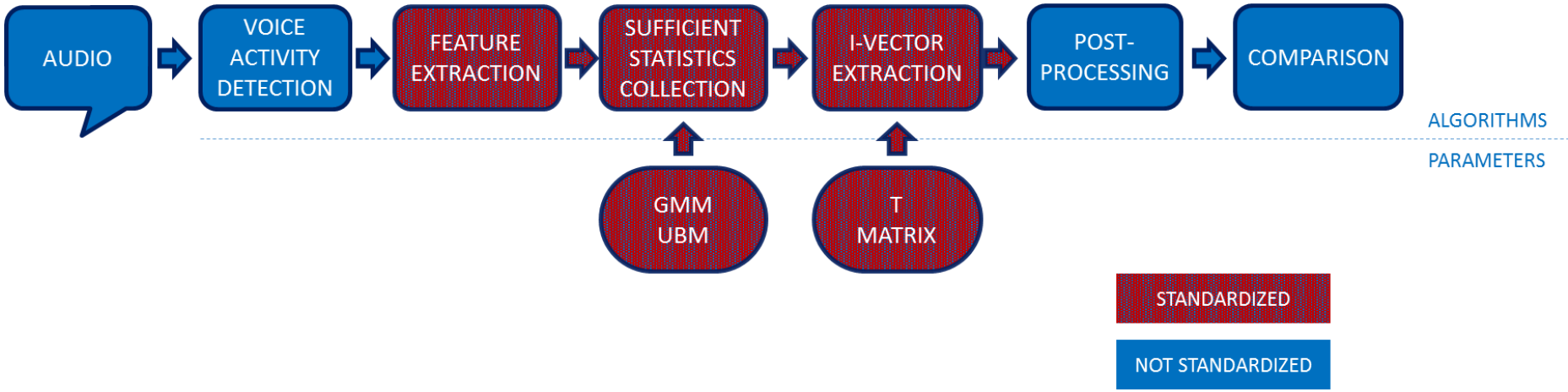


The main thing



What is needed

- Fix the core iVector extraction algorithms
- Fix the necessary parameters
- Do the necessary minimum, let people freedom to use their (own, best) VAD and scoring.
- Do it well for the core condition – **telephone**, not trying to address everything.



We WANT

- Users
  - Having interoperable systems
  - Being able to exchange speaker information without compromising content
  - within companies/agencies, across companies/agencies and across borders
- Vendors
  - Increasing the whole market (think about introduction of USB!)
- R&D labs
  - sharing iVectors between labs without lengthy discussions on configuration (not excluded though!)
  - Giving a working recipe to juniors to play with.
  - Obtaining massive data from the users

We DON'T WANT

- stop R&D (both academic and commercial) of speaker recognition technology by saying that this will be the only iVector extraction scheme forever.
  - all of us are trying to push the field further, sometimes as collaborators, sometimes as competitors.
  - We want to define a snap-shot of the best practice up to day on which we could agree.
- Earn money on licenses or patents – the proposed standard is license and patent-free
- Have something too complex and too relying on a proprietary and/or 3rd party technology.
- Present this as an ultimate forensic solution.

What is there

- http://voicebiometry.org/ - technical description, Python code with all necessary parameters (feature extraction, UBM, T-matrix)
- Google group http://groups.google.com/d/forum/voice-biometry-standard - please subscribe

Program

Honza Cernocky – intro, “why?”

Ondrej Glembek – Technical description

Petr Schwarz – Phonexia remarks

Discussion

Honza Cernocky – next steps

End 16.00, no buffet, drinks, entertainment ☹

Program

Honza Cernocky – intro, “why?”

Ondrej Glembek – Technical description

Petr Schwarz – Phonexia remarks

Discussion

Honza Cernocky – next steps

End 16.00, no buffet, drinks, entertainment ☹

Next steps

- If interested, sign-up to the google-group:
- http://groups.google.com/d/forum/voice-biometry-standard (no more personal emails).
- take the code and test it on your data
  - Report anything that you'd like to improve.
  - Please bug-fixes, not complete changes …
  - To the g-group or personally to Ondra glembek@fit.vutbr.cz
- Tell us if we can add your lab/company as supporter on the web-page.
  - Please attach a logo in reasonable resolution and a web-link.
  - You might need to consult your management.
- **Vendors: implement it to your systems**

Ext steps II.

- The real normalization (ISO/IEC, NIST, W3C …)
- Yes, but only if it has wide industrial and academic support.
- Will need help …

Thank you for your attention !

http://voicebiometry.org/