

WFBAS

Word/Phoneme-Based Audio Search

uživatelská a instalační příručka



MINISTERSTVO  
KULTURY

# Obsah

1 Úvod	2
2 Použití	2
3 Instalace	3

# 1 Úvod

Software slouží k prohledávání databáze videoarchivu a předává uživateli (MCLASS) informaci o tom, kde a s jakou pravděpodobností se hledané výrazy vyskytují. Bližší informace o prohledávání archivu lze nalézt ve článku<sup>1</sup>.

Zjednodušeně lze říci, že WFBAS prohledává pro MCLASS databázi vytvořenou SEASR-CZE nebo SEASR-ENG.

# 2 Použití

Software se spouští příkazem

```
python pixla.pyc amalach.json jsearch px01 query.json
```

Soubor `amalach.json` obsahuje konfiguraci softwaru. Definuje například, jakým způsobem bude přistupovat software k databázi. Příklad konfigurace:

```
{
  "db_uri": "mongodb://hydra3.kky.zcu.cz/amalach3/px01",
  "temp_dir": "./tmp",
  "models_dir": "./models",

  "http_port": 8002,
  "private_http_port": 8003,
  "http_processes": 4,
  "http_use_ssl": 1,
  "http_limit_results": 500,

  "ssl_certfile": "ca/amalach.zcu.cz.pem",
  "ssl_keyfile": "ca/serverkey.dec.pem",
  "ssl_ca_certs": "ca/tcs-ca-bundle.pem",

  "default_profile": "amalach_en",
```

---

<sup>1</sup>*System for fast lexical and phonetic spoken term detection in a Czech cultural heritage archive.* [dokument ve formátu PDF] dostupný z: <http://www.asmp.eurasipjournals.com/content/2011/1/10>

```

    "profiles": {
      "amalach_en": {
        "recog_model": "en/AMALACH",
        "formats": ["mp4"],
        "sample_freq": 16000
      }
    }
  }
}

```

Následuje příkaz `jsearch` který říká, že se bude vyhledávat pomocí dotazu ve formátu JSON. Argument `px01` zde upřesňuje prohledávanou oblast databáze. Poslední je samotný dotaz `query.json`. Příklad dotazu:

```

{
  "metadata": [],
  "lang": "en",
  "phrases": [{
    "words": ["HOUSE"],
    "span": 3
  }]
}

```

Odpověď softwaru je také ve formátu JSON. Pro bližší vysvětlení významu argumentů lze použít příkaz

```
python pixla.pyc --help
```

### 3 Instalace

Software se pouze kopíruje na požadované umístění.

Pro správnou funkci je potřeba nainstalovat následující balíčky Pythonu. Zřejmě nejsnazší je použití příkazu `pip`.

```

tornado
Motor
PyMongo == 2.8
shortuuid
log4mongo
isodate

```