Linking research information within the knowledge portal - the Bridge of Knowledge services for researchers

Dr Anna Wałek
Gdańsk University of Technology Library

euroCRIS Conference Dubrovnik, 11-14.05.2022
The Bridge of Knowledge

- Business
- Researchers
- Information society

- MOST WIEDZY

- open science

- publications
- manuscripts
- grants
- projects
- R&D teams
- laboratories
- research equipment
- inventions
- patents
- trademarks
- offers for business
The goal was to create **Competence Center** which provides expertise and support including trainings among scholars about different aspects of Open Science and face to face consultations and to design and build a **platform dedicated to research data** generated at the **three most important universities of Pomerania**, which will increase accessibility, coherence and reuse of science, knowledge and technology resources.

co-financed by the European Regional Development Fund
Polish National Science Centre

One of the signatories of PlanS

DMP’s in the project proposals since 2020
The Data Citation Index™ provides a single point of access to quality research data from global repositories across disciplines. Descriptive records are created for data objects and linked to literature articles in the Web of Science.™

As data citation practices increase, the Data Citation Index aims to provide a clearer picture of the full impact of research output, as well as to act as a significant tool for data attribution and discovery.

As of 2 March 2022
- Data from 443 repositories
- 12,357,621 datasets
- 1,434,939 data studies
- 264,773 software
### MOST WIEZY
your knowledge portal

#### Browse knowledge catalog

<table>
<thead>
<tr>
<th>Research Infrastructure</th>
<th>Ministry Points</th>
<th>Scientific activity</th>
<th>Repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Teams</td>
<td>Journals</td>
<td>People</td>
<td>Publications</td>
</tr>
<tr>
<td>Research Equipment</td>
<td>Publishing Houses</td>
<td>Projects</td>
<td>Repository Open Access</td>
</tr>
<tr>
<td>Laboratories</td>
<td>Conferences</td>
<td>Inventions</td>
<td>Open Research Data</td>
</tr>
</tbody>
</table>

### MOST WIEZY - WHAT IS IT?
MOST Wiedzy is a place where you can present your research. Read more about the project, create your own profile, and show your achievements.
Publication showcase

**An audio-visual corpus for multimodal automatic speech recognition**


A review of available audio-visual speech corpora and a description of a new multimodal corpus of English speech recordings is provided. The corpus containing 31 hours of recordings was created specifically to assist audio-visual speech recognition systems (AVSR) development. The database related to the corpus includes high-resolution, high-framerate stereoscopic video streams from RGB cameras, depth imaging stream utilizing Time-of-Flight...

Full text available

**Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations**

K. Lopez, J. Kasus, A. Czyżewski - *MULTIMEDIA TOOLS AND APPLICATIONS* - 2015

Evaluation of sound event detection, classification and localization of hazardous acoustic events in the presence of background noise of different types and changing intensities is presented. The methods for discerning between the events being in focus and the acoustic background are introduced. The classifier, based on a Support Vector Machine algorithm, is described. The set of features and samples used for the training of the...

Full text in external service

**Detection and localization of selected acoustic events in acoustic field for smart surveillance applications**

J. Kasus, K. Lopez, A. Czyżewski - *MULTIMEDIA TOOLS AND APPLICATIONS* - 2014

A method for automatic determination of position of chosen sound events such as speech signals and impulse sounds in 3-dimensional space is presented. The events are localized in the presence of sound reflections employing acoustic vector sensors. Human voice and impulsive sounds are detected using adaptive detectors based on modified peak-valley difference (PVA) parameter and sound pressure level. Localization based on signals...

Full text in external service
Binary Mixtures of Selected Bisphenols in the Environment: Their Toxicity in Relationship to Individual Constituents

Abstract

Bisphenol A (BPA) is one of the most popular and commonly used plasticizer in the industry. Over the past decade, new chemicals that belong to the bisphenol group have increasingly been used in industrial applications as alternatives to BPA. Nevertheless, information on the combined effects of bisphenol (BPA) analogues is insufficient. Therefore, our current study aimed to find the biological response modulations induced by the binary mixtures of BP compounds. We determined the toxicity levels in Microtox and XenoScreen YES/VAS assays for several BP analogs alone, and for their binary mixtures. The results obtained constitute the database for chemometric/biophysical models analysis to evaluate the possible interactions occurring in the mixtures. Several chemometric/biophysical models have been used (concentration addition—CA, independent action—IA and polynomial regression calculations) to realize this aim. The best fitting was found for the IA model and even in this description strong evidence for synergistic behaviors (modes of action) of some bisphenol analogue mixtures was demonstrated. Bisphenols, B and F, were proven to be of significant endocrine threat (with respect to XenoScreen YES/VAS assay); thus, their presence in mixtures (including presence in tissues of living organisms) should be most strictly monitored and reported.

Citations

Cite as

[1] CrossRef
[12] Scopus

Keywords

BISPHENOL ANALOGUES, MICROTOX®, MODEL DEVIATION RATIO, XENOSCREEN YES/VAS

Details

Category: Articles
Type: artykuł w czasopiśmie wyróżnionym w JCR
Published in: MOLECULES no. 23, pages 3226 - 3241,
ISSN: 1420-3040
Language: English
Publication year: 2018

DOI: 10.3390/molecules23123226
Assessing ecotoxicity and the endocrine potential of selected phthalates, BADGE and BFDGE derivatives in relation to environmentally detectable levels
N. Jelitowska, B. Kudlak, J. Namieśnik
2018

Bisphenols (A, S, and F) affect the basic hormonal activity determined for pharmaceuticals – Study of Saccharomyces cerevisiae
B. Kudlak, M. Włoszczyk, J. Namieśnik
2019

Modeling of pharmaceuticals mixtures toxicity with deviation ratio and best-fit functions models
M. Włoszczyk, B. Kudlak, G. Yotova
4 authors
2016

Impact of selected drugs and their binary mixtures on the germination of Sorghum bicolor (sorgo) seeds
M. Włoszczyk, B. Kudlak, J. Namieśnik
2018
Description

The MODALITY corpus is one of the multimodal database of word recordings in English. It consists of over 30 hours of multimodal recordings. The database contains high-resolution, high-framerate stereoscopic video streams and audio signals obtained from a microphone array and a laptop microphone. The corpus can be employed to develop an ASR system, as every utterance was labelled. Recordings in noisy conditions can be used to test the robustness of speech recognition systems.

The language material was based on a remote control scenario and it includes 231 words-numbers, names of months and days, a set of verbs and nouns related to a computer device control. They were read by speakers as separated words and sequences resulting in a set of 12 recording sessions per speaker. Half of the sessions were recorded in quiet conditions, the other half contained three kinds of intrusive signals (traffic, babble and factory noise).

The corpus includes recordings of 42 speakers (33 male, 9 female). The participants include 20 students and staff of Multimedia Systems Department of the Gdańsk University of Technology, 5 students of the Institute of English and American Studies of the University of Gdańsk, and 17 native English speakers.

The dataset consist of recordings and visual features for SPEAKER 33:

- sex: man
- native speaker: yes
- age: 48

The test material: SEQUENCE S2

All recordings for all speakers are available at http://www.modality.corpus.org/
Due to the size of the corpus (approx. 2.5 TB of data), every speaker’s recording was placed in a separate zip file of the size approx. 4.7 GB each.

The recordings were organized according to the speakers’ language skills. The group A (17 speakers) consists of native-speakers. Non-native speakers recordings (Polish nationals) were placed in the Group B (25 speakers).

The audio files use the Waveform Audio File Format (.wav), and contain a single PCM audio stream sampled at 44.1 kHz with 16-bit depth. The video files utilize the Matroska Multimedia Container Format (.mkv) in which a video stream in 1080p resolution, captured at 100 fps was placed after being compressed with H.264 codec (using High 4:4:4 profile).

The ‘lab’ files are text files containing the information on word positions in audio files, and follow the HTK label format. Each line of a ‘lab’ file contains the actual label preceded by start and end times (in 100 ns units) e.g.: 1239620000 1244790000 FIVE which denotes the word “five”, occurring between the 123.962 s and 124.479 s of audio.

Word-accurate SNR values calculated for every recording are also included in the ZIP file.

**Dataset file**

[SP33_SEQUENCE2.ZIP](#) 4.5 GB, MD5 efa5f7038f19504f47cb71093063d699, downloads: 0

**File details**

**License:** Custom [read]

**Details**

- **Year of publication:** 2016
- **Verification date:** 2021-07-05
- **Dataset language:** English
- **Fields of science:** Information and communication technology (Engineering and Technology), Linguistics (Humanities)
- **DOI:** [10.34608/nt1.4d97](#)
- **Series:** MODALITY corpus
- **Verified by:** Gdańsk University of Technology

**Keywords**

- language
- native speakers
- multimodal
- recordings
Core Certified Repositories

Applications are made public only once certification of a data repository has been approved by the CoreTrustSeal Board. Certification is against the version of the Core Trustworthy Data Repositories Requirements named in the link to the public application (e.g., 2017-2018). The CoreTrustSeal for Data Repositories is valid for three years from the certification date listed within the public application.
FAIR Data

Findable
Metadata and data should be findable for both humans and computers

Interoperable
Data needs to work with applications or workflows for analysis, storage and processing

Accessible
Once found, users need to know how the data can be accessed

Reusable
The goal of FAIR is to optimise data reuse via comprehensive well-described metadata
The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1. the protocol is open, free and universally implementable
A1.2. the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared and broadly applicable language for knowledge representation
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage licence
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards
Findable Accessible Interoperable Reusable ≠ Open
How do Open and FAIR intersect?

- Open
- Community benefit
- Managed data
- Self-interest
Open Data Repository – mostwiedzy.pl

- Collect and store datasets from Gdańsk Tech, GUMed and UG
- Indexing datasets in Google Data, Web of Science Data Citation Index and other services
- Technological innovations such as hosting the project on the private computing cloud and storing the data on the Ceph Object Storage.
- NoSQL database – ElasticSearch.
- Repository allows researchers to perform Big Data Analysis by the Apache Zeppelin GUI on the supercomputer Tryton (40,000 cores, 1.5 PFLOPS).
Open Science Competence Center

• Assistance and on-site tailoring training among researchers from all scientific disciplines that include Data Management Plan, open licensing or metadata standards

• Workshops regarding different aspects of Open Science as well as scholarly communication

• InfoKit regarding Open Research Data

• Metadata support

• Journals & Conference Proceedings indexing and publishing support

• Evaluation & Bibliometric support
BRIDGE OF DATA

Open Research Data

Project "MOST DANYCH"
is co-financed by the European Regional Development Fundwithin the Operational Programme Digital Poland 2014-2020
Thank you