

PerSCiDO_Grenoble_Alpes :

Principes et fonctionnalités d'une plateforme ouverte et interopérable de partage de jeux de données

<https://persyval-platform.imag.fr/perscido/web/>

Responsable: Marie-Christine ROUSSET

Lucie Albaret⁽⁵⁾, Brigitte Bidegaray^(1,2,4), Pierre Hébert⁽²⁾,
Fabrice Jouanot⁽³⁾, Alireza Moussaei⁽¹⁾

Université Grenoble Alpes

(1) CNRS, (2) Labex PERSYVAL-lab, (3) LIG, (4) LJK,

(5) Service inter-établissement de la Documentation Université Grenoble Alpes - Grenoble INP



Objectifs

- Construire une plateforme de partage de jeux de données de recherche
 - Découplant le problème de stockage des jeux de données de leur description
 - Centrée sur des méta-données riches et flexibles
 - Permettant une interrogation avancée de ces méta-données
 - Facilement interopérable avec d'autres plateformes

Principes

- Suivre les standards
 - du Linked Open Data en termes de modèle de données (RDF) et aussi de vocabulaires spécialisés de métadonnées comme Dublin Core, Friend Of a Friend, Creative Commons, etc
 - émergents de consortiums internationaux comme DataCite, FaBio, Radar, etc ... qui visent la définition de standards pour citer et décrire des données de recherche.
- ⇒ **Cohabitation dans une même base de données RDF de différents vocabulaires spécialisés correspondant à différents espaces de noms identifiés par des préfixes raccourcis**
- Anticiper les usages et inciter aux bonnes pratiques

Inciter aux bonnes pratiques

- Pousser les chercheurs à référencer leurs jeux de données par des **identifiants externes persistants** (HAL, DOI, etc ...)
 - Un DOI (Digital Object Identifier), prôné par DataCite
 - => Convention signée avec l'INIST (représentant français de DataCite) pour que PERSYVAL-lab puisse délivrer des DOIs
- Pousser les chercheurs à anticiper **la citation** souhaitée pour leur jeu de données
 - => champ pré-rempli en cas d'un DOI existant
- Pousser les chercheurs à **préciser le droit d'usage** de leurs jeux de données **par une licence Creative Commons**
 - => Menu déroulant avec les différentes licences fournies sous la forme d'un vocabulaire contrôlé
- **Eviter tant que possible la saisie de chaînes de caractères « libres »** pour remplir les valeurs de champs à renseigner
 - => Menus déroulants avec des valeurs prédéfinies (des constantes dans la BD)

Moyens humains et méthodes de travail

- Développement
 - Ali Moussaei (15 mois + 12 mois CDD ingénieur CNRS)
 - Pierre Hébert (à 80% sur PerSCiDO): ingénieur Web sur fonds propres du labex
- Spécifications et suivi
 - Groupe de travail constitué de futurs usagers (chercheurs, ingénieurs de recherche, enseignants chercheurs)
 - ⇒ Identification de plusieurs types de jeux de données et des méta-données utiles associées
 - ⇒ Identification de méta-données communes à tous les jeux de données et de méta-données spécifiques selon le type de données
 - Encadrement et suivi de la mise en œuvre: L. Albaret, B. Bidegaray, F. Jouanot, M-C Rousset
- Jalons réguliers
 - **PerSCiDO v1:**
 - Janvier- Aout 2016: **maquette** de l'interface et des fonctionnalités de **dépôt** et de **recherche avancée** de jeux de données
 - Septembre 2017: Ouverture de la plateforme

Demo

Version 1.5 (26.05.2016)





PerSCiDO facilitates the exploration of research datasets.
Share your research datasets using PerSCiDO !

Numbers

Datasets :5
Downloaded : 32585
Publications : 2587

Explore PerSCiDO research data collections and related publications

Keywords...

Recent datasets

16 05 19

From Medical Images to Computational Medicine

Nicholas CAGE

This inventory contains a set of terms that are relevant to the study of medical history. The inventory is organised as a set of "heading terms", belonging to one of seven different semantic categories, each of which is accompanied by a set of semantically-related terms. There are around 175,000 heading terms. The nature of the semantic relationship holding between the heading term and each related term varies. A pair of related terms may, for example be synonyms of each other, one term may be more or less specific than the other, one may be a part of the other, one may be used in the treatment of the other, etc. The unique feature of our terminological inventory is that the sematically-related terms may correspond to terms used within different periods of time, and which may not be in common usage today.



16 05 19

X-ray diffraction images for DPF3 tandem PHD fingers co-crystallized with an acetylated histone-derived peptide

Tempel Wolfram

This submission includes a tar archive of bziped diffraction images recorded with the ADSC Q315r detector at the Advanced Photon Source of Argonne National Laboratory, Structural Biology Center beam line 19-ID. Relevant meta data can be found in the ...



16 05 19

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

Moscato Pablo

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification



News

Bonne année à tous !!!
#HappyNewYear
#2016
21 04 2016

Data QA/QC: a natural history perspective. Free #DWS16 Webinar
Tues Jan 12th
9am Pacific
21 04 2016

Requesting data isn't harassment, and refusing to share data isn't science
<http://www.statnews.com/2015/12/23/sharing-data-science/>
21 04 2016

A natural history perspective. Free #DWS16 Webinar
Tues Jan 31th
9am Atlantic
21 04 2016

soumission d'un nouveau dataset

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO. Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Let your dataset easily citeable and discoverable

If your dataset does not have a permanent identifier, we can provide a DOI (Digital Object Identifier) for it.

The [DOI system](#) is an international standard recommended to identify and link any digital resource (publication, dataset,...). We encourage you to ask us a DOI for your dataset even if it has already another type of identifier. More information can be found on [DataCite](#).

Your dataset has already an identifier

Your dataset does not have any identifier

soumission d'un nouveau dataset

[1. General Information](#)[2. Content Description](#)[3. Datatype Content](#)[4. Data Access](#)

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO.
Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Let your dataset easily citeable and discoverable

If your dataset does not have a permanent identifier, we can provide a DOI (Digital Object Identifier) for it.

The [DOI system](#) is an international standard recommended to identify and link any digital resource (publication, dataset,...). We encourage you to ask us a DOI for your dataset even if it has already another type of identifier. More information can be found on [DataCite](#).


Do you want to ask for a Doi (Digital Object Identifier)? you can ask for it now or later.

Yes No

Your DOI will be generated after validation by the committee

[Next](#)

soumission d'un nouveau dataset





1. General Information

2. Content Description

3. Datatype Content


4. Data Access


Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO.
Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Depositor:	<input type="text" value="Marie-Christine"/>	<input type="text" value="Rousset"/>	
Depositor's identifier (if any):	<input type="text" value="your Orcid id"/>	<input type="text" value="your ArXiv id"/>	<input type="text" value="your Hal id"/>
Institution/Laboratory: *	<input type="text" value="Laboratoire d'Informatique de Grenoble"/>		
Access right: *	<input checked="" type="radio"/> Open <input type="radio"/> Restricted <small>Permission to upload your dataset (Open: everyone, Restricted: after your permission)</small>		
License:	<input type="text" value="CC_BY-NC_4.0_Attribution_NonCommercial"/>		

The choice of your license is under your responsibility

[PowerdBy ...](#) - [Term of service](#) - [Contact us](#)





soumission d'un nouveau dataset

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCIDO.
Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Title of your dataset: *

Description: *

Keywords:*

Subject: *

<input type="checkbox"/> Agriculture	<input type="checkbox"/> Chemistry	<input type="checkbox"/> History
<input type="checkbox"/> Architecture	<input checked="" type="checkbox"/> Computer Science	<input type="checkbox"/> Information Technology
<input type="checkbox"/> Arts and Media	<input type="checkbox"/> Economics	<input type="checkbox"/> Mathematics
<input type="checkbox"/> Astrophysics and Astronomy	<input type="checkbox"/> Engineering	<input type="checkbox"/> Medicine
<input type="checkbox"/> Behavioural Sciences	<input type="checkbox"/> Environmental Science and Ecology	<input type="checkbox"/> Physics
<input type="checkbox"/> Biochemistry	<input type="checkbox"/> Geography	<input type="checkbox"/> Social Sciences
<input type="checkbox"/> Biology	<input type="checkbox"/> Geological Sciences	<input type="checkbox"/> Other

Related publications to your dataset:

Next

soumission d'un nouveau dataset

PERSYVAL - Lab

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO. Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Please select the data type of your dataset

Data type

- Simulation data ?
- Survey data ?
- Textual data ?
- Trace data ?
- Video data ?
- Web data ?

Automatic task,

<input type="checkbox"/> Classification	<input type="checkbox"/> Pattern extraction	<input type="checkbox"/> Regression Analysis
<input type="checkbox"/> Clustering	<input type="checkbox"/> Prediction	<input type="checkbox"/> Rule extraction
<input type="checkbox"/> Dimension Reduction	<input type="checkbox"/> Preference learning	<input type="checkbox"/> Visualisation

Next

soumission d'un nouveau dataset

PERSYVAL - Lab
FINANCE PAR ANR INVESTISSEMENTS D'AVENIR

1. General Information
2. Content Description
3. Datatype Content
4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO.
Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Please select the data type of your dataset

Trace data

Execution Trace

- Interaction Trace ?
- Execution Trace ?
- Sensor Trace ?
- Systems Trace ?
- Other Trace ?

If your dataset has been processed for an automa

please select the corresponding task(s) below

Anomaly detection
 Classification
 Clustering
 Dimension Reduction

Grammatical inference
 Pattern extraction
 Prediction
 Preference learning

Regression Analysis
 Rule extraction
 Visualisation

Next

soumission d'un nouveau dataset

PERSYVAL - Lab
ANR

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO.
Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Please select the data type of your dataset

Trace data

Execution Trace

If your dataset has been processed for an automatic task,

please select the corresponding task(s) below

<input type="checkbox"/> Anomaly detection	<input type="checkbox"/> Grammatical inference	<input type="checkbox"/> Regression Analysis
<input type="checkbox"/> Classification	<input type="checkbox"/> Pattern extraction	<input type="checkbox"/> Rule extraction
<input type="checkbox"/> Clustering	<input type="checkbox"/> Prediction	<input type="checkbox"/> Visualisation
<input type="checkbox"/> Dimension Reduction	<input type="checkbox"/> Preference learning	

Next

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO. Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Informations about your dataset

Size:

- Less than 10 Mb
 Less than 100 Mb
 Less than 500 Mb
 Less than 4 Gb
 More than 4 Gb

To facilitate the use of your dataset, we encourage you to give indications about its size

Encoding data format

CSV

To facilitate the use of your dataset, we encourage you to give indications about its format

Citation for your dataset:

To facilitate the citation of your dataset we encourage you to provide the recommended text for citing it

Illustrative logo:

Browse... logo.gif

Read-me file for your dataset:

Browse... kptrace.meta.readme

Data storage mode:

You can either upload your dataset for storage on the PerSCiDO platform, or you can provide the URL where your dataset can be found.

- Upload your dataset for storage on Percido platform
 Provide the url address of your dataset:

Archive file name:

Browse... tsrec_TF1_nomodif_CRL.zip

UploadFile

soumission d'un nouveau dataset

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCiDO. Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Informations about your dataset

Size: Less than 10 Mb Less than 100 Mb Less than 500 Mb Less than 4 Gb More than 4 Gb
To facilitate the use of your dataset, we encourage you to give indications about its size

Encoding data format
To facilitate the use of your dataset, we encourage you to give indications about its format

Citation for your dataset:
To facilitate the citation of your dataset we encourage you to provide the recommended text for citing it

Illustrative logo: logo.gif

Read-me file for your dataset: kptrace.meta.readme

Data storage mode:
You can either upload your dataset for storage on the PerSCiDO platform, or you can provide the URL where your dataset can be found.

The upload of your archive file is complete.

Finish

soumission d'un nouveau dataset

1. General Information

2. Content Description

3. Datatype Content

4. Data Access

Please describe your dataset in as much detail as possible. A detailed description will make it easier for others to find your dataset in PerSCIDO. Fields marked with an asterisk (*) are required. For more information on expected content for a field, mouse over the ? icon.

Details of your Dataset

Creation date : 16 05 20
Title : Trace d'execution audio video
Description : Desynchronisation audio video sample TF1 (board STMicroelectronics) scenario TSrecord format KPTRAC
Contributor : Marie-Christine Rousset
File readme : kptrace.meta.readme
Fileurl :
Filezip : tsrec_TF1_nomodif_CRL.zip
Format : csv
Size : 500
Accessright : Open
Datatype : Trace data
Subdatatype : Execution Trace
Institution : Laboratoire d'Informatique de Grenoble
License : CC_BY-NC_4.0_Attribution_NonCommercial
keywords : Execution Trace Audio Video
subjects : Computer Science



Validation du dépôt

- Par un comité « éditorial »
 - Qui peut demander de renseigner certains champs de méta-données non renseignés
 - Qui s'assure que les données ne contiennent pas d'informations personnelles ou sensibles
 - Qui valide de nouvelles valeurs et/ou de nouveaux champs de méta-données
 - ajoutées à l'ontologie

Behind the scene: une ontologie

```
foaf:firstName
  rdf:type owl:DatatypeProperty ;
  rdfs:domain foaf:Person ;
  rdfs:range xsd:string ;
```

```
.....
dcterms:description
  rdf:type owl:DatatypeProperty ;
  rdfs:domain fabio:Dataset ;
  rdfs:range xsd:string ;
```

```
.....
dcterms:subject
  rdf:type owl:ObjectProperty ;
  rdfs:comment "Property relating the local identifier of a dataset to the scientific field it is related to"@en ;
  rdfs:comment "Propriete reliant l'identifiant local d'un jeu de donnees a la discipline scientifique a laquelle il se rapporte"@fr ;
  rdfs:domain fabio:Dataset ;
  rdfs:label "Scientific Field"@en ;
  rdfs:label "Discipline Scientifique"@fr ;
  rdfs:range :Subject_Area ;
```

```
.....
:Subject_Area
  rdf:type owl:Class ;
  rdfs:comment "Class denoting the possible scientific fields of datasets"@en ;
  rdfs:comment "Classe regroupant les differentes disciplines scientifiques auxquelles peuvent se rattacher un jeu de donnees"@fr ;
  rdfs:label "Discipline Scientifique"@fr ;
  rdfs:label "Scientific Field"@en ;
  rdfs:subClassOf :toValidate ;
```

```
.....
radar:Chemistry
  rdf:type :Subject_Area ;
  rdfs:label "Chemistry"@en ;
  rdfs:label "Chimie"@fr ;
  :isValid true;
```

```
.....
radar:Computer_Science
  rdf:type :Subject_Area ;
  rdfs:label "Computer Science"@en ;
  rdfs:label "Informatique"@fr ;
  :isValid true;
```


... et des méta-données associées à des datasets

```
@prefix cc: <http://creativecommons.org/licences/> .
@prefix : <http://example.org/perscido#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix fabio: <https://svn.code.sf.net/p/semublishing/code/FaBiO/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix datacite: <http://purl.org/spar/datacite/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix dcmitype: <http://purl.org/dc/dcmitype/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix radar: <http://www.radar-projekt.org/display/> .
@prefix swpo: <http://sw-portal.deri.org/ontologies/swportal#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dcterms: <http://dublincore.org/documents/dcmi-terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
```

```
:DS6 a fabio:Dataset ;
dcterms:contributor :user14 ;
dcterms:description "MovieLens 1M dataset enriched with IMDB on movie attributes" ;
dcterms:subject radar:Computer_Science ;
dcterms:title "MovieLens+IMDb" ;
:AccessRight :Open ;
:Corresponding_tasks :CT_Clustering , :CT_Prediction , :CT_Visualisation ;
:DownloadInfo :Download2 , :Download3 , :Download8 ;
:HasData_format :DF_csv ;
:HasKeywords :KW_RATING , :KW_WEB , :KW_SOCIAL , :KW_COLLABORATIVE , :KW_USER ;
:Has_Citation "test" ;
:Has_Logo_Url <https://persyval-platform.imag.fr/perscido/web/datasetFiles/DS6/logo.png> ;
:Has_ReadMeFile_Url <https://persyval-platform.imag.fr/perscido/web/datasetFiles/DS6/README.txt> ;
:Has_ZipFile_Url <https://persyval-platform.imag.fr/perscido/web/datasetFiles/DS6/movielens-imdb.zip> ;
:hasDataType :Donnees_Web ;
:hasInstitution :IN_OHIO_STATE_UNIVERSITY ;
:hasState :validated ;
:has_max_size 100 ;
:has_min_size 10 ;
:lastStageDepot :f4 ;
:questionDoi true ;
:startDepositDate "2016-10-13T17:42:05+01:00"^^xsd:dateTime ;
:validatedDate "2016-10-28T17:42:05+01:00"^^xsd:dateTime .
```

```
:user14 a foaf:Person ;
:hasAccount true ;
foaf:firstName "Behrooz OMIDVAR" ;
foaf:lastName "behrooz" ;
foaf:mbox <mailto:behrooz.omidvar@gmail.com>
```

```
:KW_RATING a :Keywords ;
rdfs:label "rating"@en ;
:isValid true .
```

```
:KW_WEB a :Keywords ;
rdfs:label "web"@en ;
:isValid true .
```

```
:KW_SOCIAL a :Keywords ;
rdfs:label "social"@en ;
:isValid true .
```

```
:KW_COLLABORATIVE a :Keywords ;
rdfs:label "collaborative"@en ;
:isValid true .
```

```
:KW_USER a :Keywords ;
rdfs:label "user"@en ;
:isValid true .
```

```
:IN_OHIO_STATE_UNIVERSITY
a :Institution ;
rdfs:label "Ohio State University"@en ;
:isValid true .
```

Recherche de datasets

PerSCiDO facilitates the exploration of research datasets.
Share your research datasets using PerSCiDO !

Numbers
Datasets : 3
Downloaded : 32585
Publications : 2587

Explore PerSCiDO research data collections and related publications

Keywords...

Recent datasets

16 05 19

From Medical Images to Computational Medicine
Nicholas CAGE
This inventory contains a set of terms that are relevant to the study of medical history. The inventory is organised as a set of "heading terms", belonging to one of seven different semantic categories, each of which is accompanied by a set of semantically-related terms. There are around 175,000 heading terms.



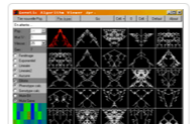
16 05 19

X-ray diffraction images for DPF3 tandem PHD fingers co-crystallized with an acetylated histone-derived peptide
Tempel Wolfram
This submission includes a tar archive of bzipped diffraction images recorded with the ADSC Q315r detector at the Advanced Photon Source of Argonne National Laboratory, Structural Biology Center beam line 19-ID. Relevant meta data can be found in the ...



16 05 19

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"
Moscato Pablo
Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...



News

Presentation of PerSCiDO at data4ist day (Paris, May 23 2016)
20 05 2016

Data QA/QC: a natural history perspective. Free #DWS16 Webinar Tues Jan 12th 9am Pacific
21 04 2016

Requesting data isn't harassment, and refusing to share data isn't science
<http://www.statnews.com/2015/12/23/sharing-data-science/>
21 04 2016

Recherche de datasets

PerSCiDO facilitates the exploration of research datasets.
Share your research datasets using PerSCiDO !

Numbers
Datasets :3
Downloaded : 32585
Publications : 2587

Explore PerSCiDO research data collections and related publications

medic

medicine
medic
medical

search

Submit a new dataset

Recent datasets

Recently Published | By Subject | By Data types | Size 4 Gb

16 05 19 Restricted Trace data

From Medical Images to Computational Medicine

Nicholas CAGE

This inventory contains a set of terms that are relevant to the study of medical history. The inventory is organised as a set of "heading terms", belonging to one of seven different semantic categories, each of which is accompanied by a set of semantically-related terms. There are around 175,000 heading terms.

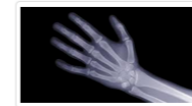


16 05 19 Open Trace data

X-ray diffraction images for DPF3 tandem PHD fingers co-crystallized with an acetylated histone-derived peptide

Tempel Wolfram

This submission includes a tar archive of bzipped diffraction images recorded with the ADSC Q315r detector at the Advanced Photon Source of Argonne National Laboratory, Structural Biology Center beam line 19-ID. Relevant meta data can be found in the ...

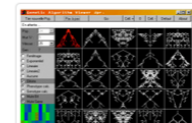


16 05 19 Restricted Survey data

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

Moscato Pablo

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...



News

Presentation of PerSCiDO at data4ist day (Paris, May 23 2016)
20 05 2016

Data QA/QC: a natural history perspective. Free #DWS16 Webinar Tues Jan 12th 9am Pacific
21 04 2016

Requesting data isn't harassment, and refusing to share data isn't science
<http://www.statnews.com/2015/12/23/sharing-data-science/>
21 04 2016

Recherche de datasets

Search

Refine

▼ By Subject

- Agriculture
- Architecture
- Arts and Media
- Astrophysics and Astronomy
- Behavioural Sciences
- Biochemistry
- Biology
- Chemistry
- Computer Science
- Economics
- Engineering
- Environmental Science and Ecology
- Geography
- Geological Sciences
- History
- Information Technology
- Mathematics
- Medicine
- Medicine
- Other
- Physics
- Social Sciences

► By year

► By DataType

► By Access

Search results

3 results Date ▾

Keywords Subject Year Datatypes Access Right

16 05 19
Restricted
Trace data

From Medical Images to Computational Medicine

Nicholas CAGE

This inventory contains a set of terms that are relevant to the study of medical history. The inventory is organised as a set of "heading terms", belonging to one of seven different semantic categories, each of which is accompanied by a set of semantically-related terms. There are around 175,000 heading terms.

16 05 19
Open
Trace data

X-ray diffraction images for DPF3 tandem PHD fingers co-crystallized with an acetylated histone-derived peptide

Tempel Wolfram

This submission includes a tar archive of bziped diffraction images recorded with the ADSC Q315r detector at the Advanced Photon Source of Argonne National Laboratory, Structural Biology Center beam line 19-ID. Relevant meta data can be found in the ...

16 05 19
Restricted
Survey data

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

Moscato Pablo

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...

Recherche de datasets

Search

Refine

▼ By Subject

- Agriculture
- Architecture
- Arts and Media
- Astrophysics and Astronomy
- Behavioural Sciences
- Biochemistry
- Biology
- Chemistry
- Computer Science
- Economics
- Engineering
- Environmental Science and Ecology
- Geography
- Geological Sciences
- History
- Information Technology
- Mathematics
- Medicine
- Medicine
- Other
- Physics
- Social Sciences

► By year

► By DataType

► By Access

Search results

2 results Date ▾


Keywords Subject Year Datatypes Access Right

16 05 19
Restricted
Trace data

From Medical Images to Computational Medicine

Nicholas CAGE

This inventory contains a set of terms that are relevant to the study of medical history. The inventory is organised as a set of "heading terms", belonging to one of seven different semantic categories, each of which is accompanied by a set of semantically-related terms. There are around 175,000 heading terms.

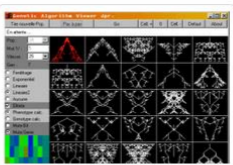


16 05 19
Restricted
Survey data

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

Moscato Pablo

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...



Recherche de datasets

Search

Refine

▸ By Subject

▸ By year

▾ By DataType

- Experimental data
- Image data
- Instrumentation data
- Simulation data
- Survey data
- Textual data
- Trace data
- Video data
- Web data
- Other

▸ By Access

Search results

2 results Date ▾

Keywords
Subject
Year
Datatypes
Access Right

16 05 19 Restricted Trace data

From Medical Images to Computational Medicine

[Nicholas CAGE](#)

This inventory contains a set of terms that are relevant to the study of medical history. The inventory is organised as a set of "heading terms", belonging to one of seven different semantic categories, each of which is accompanied by a set of semantically-related terms. There are around 175,000 heading terms.

16 05 19 Restricted Survey data

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

[Moscato Pablo](#)

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...

Recherche de datasets

Search

Refine

▸ By Subject

▸ By year

▾ By DataType

- Experimental data
- Image data
- Instrumentation data
- Simulation data
- Survey data
- Textual data
- Trace data
- Video data
- Web data
- Other

▸ By Access

Search results

1 results Date ▾

Keywords
Subject
Year
Datatypes
Access Right

16 05 19

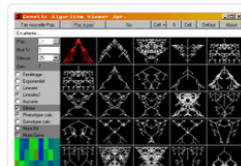
Restricted

Survey data

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

[Moscato Pablo](#)

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...



Recherche de datasets

Search

Refine

- ▶ By Subject
- ▶ By year
- ▶ By DataType
- ▼ By Access
 - Restricted
 - Open

Search results

1 results Date ▾

Keywords
Subject
Year
Datatypes
Access Right

16 05 19

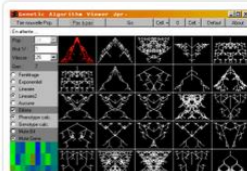
Restricted

Survey data

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

[Moscato Pablo](#)

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...



Recherche de datasets

←
16 05 19
Survey data
Restricted

Supporting datasets PubFig05 for: "Heterogeneous Ensemble Combination Search using Genetic Algorithm for Class Imbalanced Data Classification"

Contributor: [Moscato Pablo](#)

Faces Dataset: PubFig05 This is a subset of the "PubFig83" dataset [1] which provides 100 images each of 5 most difficult celebrities to recognise (referred as class in the classification problem). For each celebrity persons, we took 100 images and ...

ReadMeFile i

Readme.txt	15 02 16	152 Ko
----------------------------	----------	--------

Files ☰

Files	Date	Size
+ File.zip	25 02 16	4 Mo 🔒 Download

RealtedPublication 👁

By : [Nicholas Ayache](#), [Laurent D. Cohen](#) and [Isaac Cohen](#)
 Title : Using deformable surfaces to segment 3-D images and infer differential structures
 Date : 1992
 Available on : <http://basepub.dauphine.fr/xmlui/handle/123456789/6872>

Publication date : 16 05 19
identifier
Insitution/Laboratory : IN3
identifier :
Subjects : SI1
Keywords :

Cite as
 CI3
Export as
[BibTeX](#), [DataCite](#), [XML](#), [EndNote](#)

Metrics
 Record views 214
 Document Downloads 279

Behind the scene: le langage de requêtes SPARQL

- Trouver les datasets et leur titre, déposés par un chercheur du LIG
- Trouver les datasets de taille inférieure à 1Go en libre accès et en format csv
- Trouver des données de type « enquête » qui sont associées à des publications dans le domaine de la santé.

Travail en cours et évolution

- Décrire via PerSciDO, et stocker sur SUMMER, de très gros jeux de données
 - En simplifiant (pour l'utilisateur) le dépôt et le téléchargement tout en préservant la sécurité de l'ensemble du système
- Etendre la plateforme à la description d'algorithmes, de protocoles expérimentaux, de chaînes de traitement de données
 - Nouvelles méta-données
 - Liage entre jeux de données