CaosDB - An open scientific database

Alexander Schlemmer, Henrik tom Wörden, Timm Fitschen, Daniel Hornung, Ulrich Parlitz, and Stefan Luther

Research Group Biomedical Physics, MPI for Dynamics and Self-Organization, Göttingen

IndiScale GmbH

2019-09-23

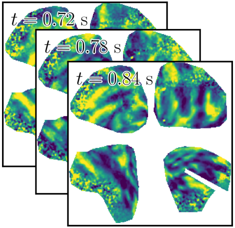
Caosdb
an open scientific database

Current Challenges in Scientific Data Management

Current Challenges in Scientific Data Management

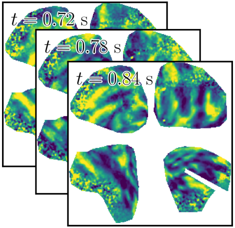
Diverse and Complex Data

Current Challenges in Scientific Data Management

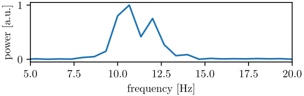


Diverse and Complex Data

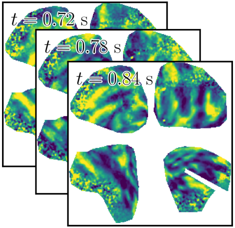
Current Challenges in Scientific Data Management



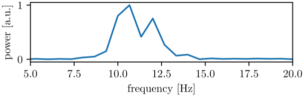
Diverse and Complex Data

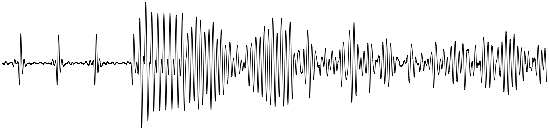


Current Challenges in Scientific Data Management

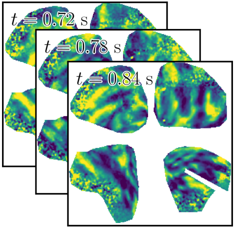


Diverse and Complex Data

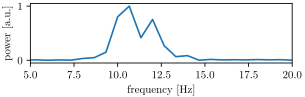


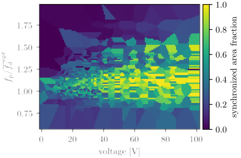


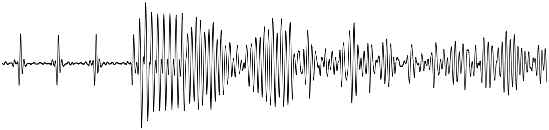
Current Challenges in Scientific Data Management



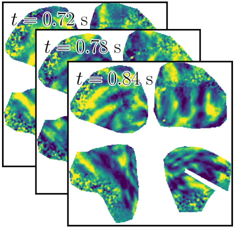
Diverse and Complex Data



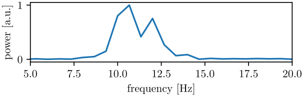


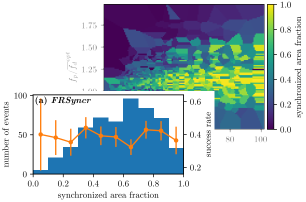


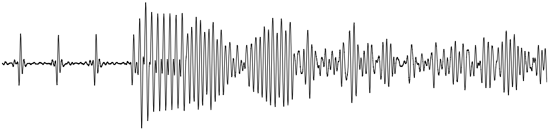
Current Challenges in Scientific Data Management



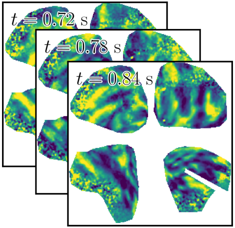
Diverse and Complex Data

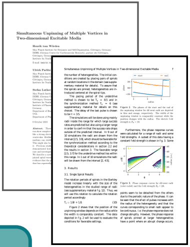






Current Challenges in Scientific Data Management





Simultaneous Unpinning of Multiple Vortices in Two-dimensional Excitable Media

Henrik tom Wörden

Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany

DZHK (German Center for Cardiovascular Research), partner site Göttingen,

Ulrich Parlitz

Stefan Luther

9 October 2018

Abstract.

Simultaneous Unpinning of Multiple Vortices in Two-dimensional Excitable Media 7

the number of heterogeneities. The initial conditions are created by placing pairs of spirals at random locations in the domain (see supplementary material for details). To assure that the spirals are pinned, heterogeneities are introduced centered at the spiral tips.

The pacing period of the underdrive method is chosen to be $T_p = 8.5$ and in the synchronization method $T_p = 6$ (see supplementary material for details on this choice). The delay of the last pulse is chosen to be $\tau = 3.5$.

The simulations will be done using mainly radii inside the range for which large success rates are predicted but also using a larger range in order to confirm that the success rate drops outside of the predicted interval. In 9 out of 10 simulations the radii are drawn from the interval [2.5, 3.9], which should be favorable for the synchronization method according to the theoretical considerations in section 2.2 and the results in section 3. The favorable range [2.5, 3.7] for the underdrive method lies within this range. In 1 out of 10 simulations the radii will be drawn from the interval [2, 4.5].

3. Results

3.1. Single Spiral Results

The rotation periods of spirals in the Barkley model increase linearly with the size of the heterogeneities in the studied range of radii (see supplementary material fig. 12). Thus, we will use this relation to calculate the rotation period accordingly

$$T_s = 1.8r + 1.8. \quad (10)$$

Figure 2 shows that the position of the unpinning window depends on the radius while the width is comparably constant. The data depicted in fig. 2 will be used to evaluate the conditions for favorable settings.

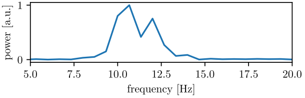
Figure 2. The phases of the start and the end of the unpinning window for different radii are depicted in blue and orange, respectively. The width of the unpinning window is comparably constant while the position changes with the radius. The electric field strength is $E_0 = 10$.

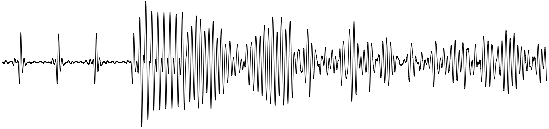
Furthermore, the phase response curves were calculated for a range of radii and some field strengths. A subset of this data having a constant field strength is shown in fig. 3. Some

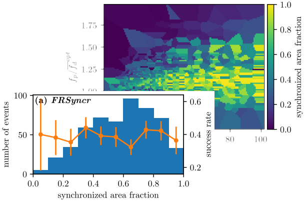
Figure 3. Phase response curves for different radii (color coded) and the field strength $E_0 = 26$.

points seem to be detached from the others due to the periodic nature of the phase. It can be seen that the effect of pulses increases with the radius of the heterogeneity and that the curves corresponding to small radii appear to be continuous. I.e the phase response does not change abruptly. However, the phase response of spirals pinned at larger heterogeneities have a point where an abrupt change occurs.

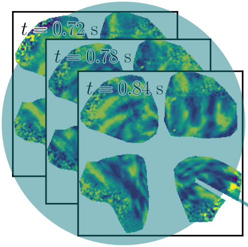
Diverse and Complex Data

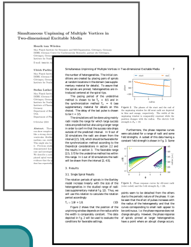






Current Challenges in Scientific Data Management





Simultaneous Unpinning of Multiple Vortices in Two-dimensional Excitable Media

Henrik tom Wörden

Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
DZHK (German Center for Cardiovascular Research), partner site Göttingen,
Göttingen, Germa
Institute for Nonli
E-mail: henrik.to

Ulrich Parlitz

Max Planck Instit
DZHK (German C
Göttingen, Germa
Institute for Nonli

Stefan Luther

Max Planck Instit
DZHK (German C
Göttingen, Germa
Institute for Nonli
Institute of Pharma
Germany.
Department of Bio
USA.
Department of Phy

9 October 2018

Abstract. Ther
can show complex
like a strong electr
ventricular fibrillat
medium can stabil
This might also be
is. Previous studi
demonstrated how
case and investigat
We describe a nove
pinned spiral wave
evidence that the p
that has reportedly

Simultaneous Unpinning of Multiple Vortices in Two-dimensional Excitable Media 7

the number of heterogeneities. The initial conditions are created by placing pairs of spirals at random locations in the domain (see supplementary material for details). To assure that the spirals are pinned, heterogeneities are introduced centered at the spiral tips.

The pacing period of the underdrive method is chosen to be $T_p = 8.5$ and in the synchronization method $T_p = 6$ (see supplementary material for details on this choice). The delay of the last pulse is chosen to be $\tau = 3.5$.

The simulations will be done using mainly radii inside the range for which large success rates are predicted but also using a larger range in order to confirm that the success rate drops outside of the predicted interval. In 9 out of 10 simulations the radii are drawn from the interval [2.5, 3.9], which should be favorable for the synchronization method according to the theoretical considerations in section 2.2 and the results in section 3. The favorable range [2.5, 3.7] for the underdrive method lies within this range. In 1 out of 10 simulations the radii will be drawn from the interval [2, 4.5].

3. Results

3.1. Single Spiral Results

The rotation periods of spirals in the Barkley model increase linearly with the size of the heterogeneities in the studied range of radii (see supplementary material fig. 12). Thus, we will use this relation to calculate the rotation period accordingly

$$T_s = 1.8r + 1.8. \quad (10)$$

Figure 2 shows that the position of the unpinning window depends on the radius while the width is comparably constant. The data depicted in fig. 2 will be used to evaluate the conditions for favorable settings.

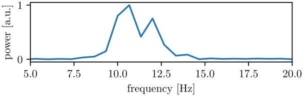
Figure 2. The phases of the start and the end of the unpinning window for different radii are depicted in blue and orange, respectively. The width of the unpinning window is comparably constant while the position changes with the radius. The electric field strength is $E_0 = 10$.

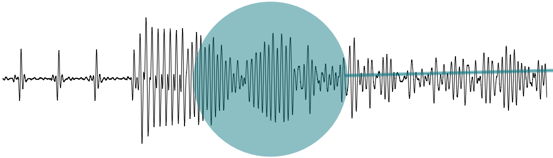
Furthermore, the phase response curves were calculated for a range of radii and some field strengths. A subset of this data having a constant field strength is shown in fig. 3. Some

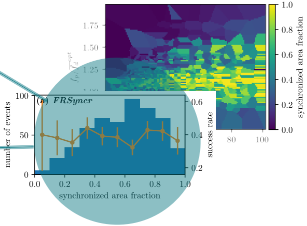
Figure 3. Phase response curves for different radii (color coded) and the field strength $E_0 = 26$.

points seem to be detached from the others due to the periodic nature of the phase. It can be seen that the effect of pulses increases with the radius of the heterogeneity and that the curves corresponding to small radii appear to be continuous. I.e the phase response does not change abruptly. However, the phase response of spirals pinned at larger heterogeneities have a point where an abrupt change occurs.

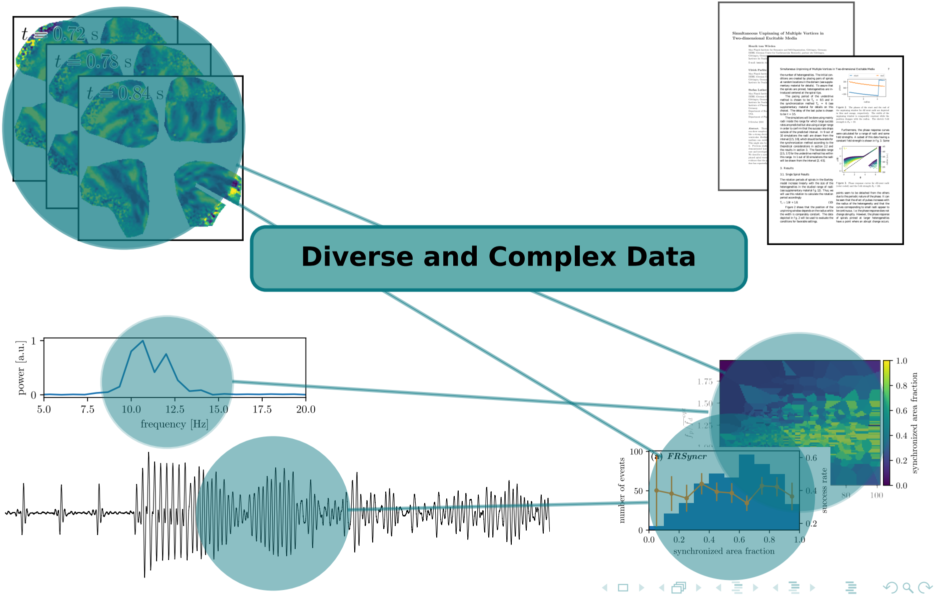
Diverse and Complex Data





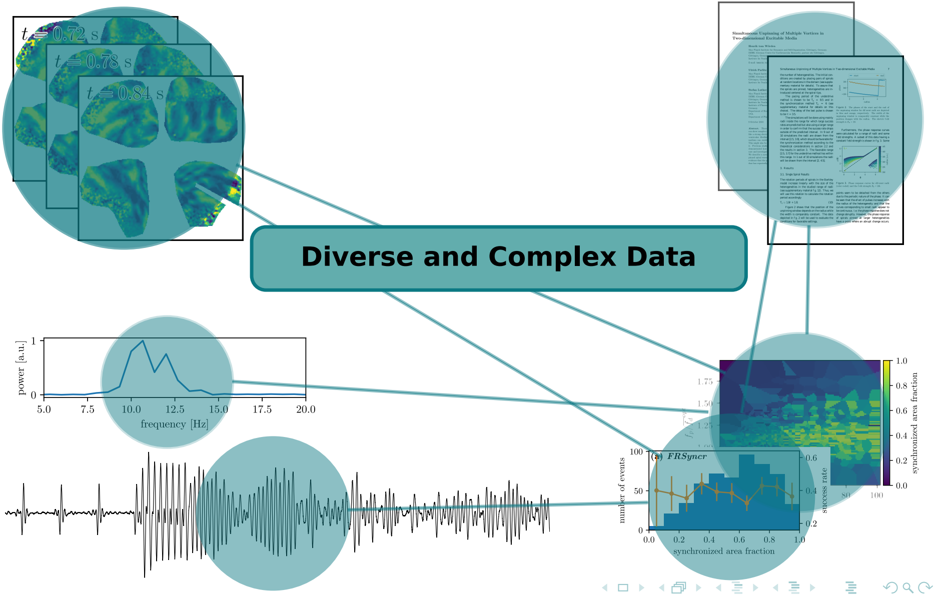


Current Challenges in Scientific Data Management



Diverse and Complex Data

Current Challenges in Scientific Data Management



Diverse and Complex Data

Simultaneous Unpinning of Multiple Vortices in Two-dimensional Excitable Media

Challenges in Scientific Data Management

- Standardization of file storage

Challenges in Scientific Data Management

- Standardization of file storage
- Data formats

Challenges in Scientific Data Management

- Standardization of file storage
- Data formats
- Metadata

Challenges in Scientific Data Management

- Standardization of file storage
- Data formats
- Metadata
- Linking of raw data, processed data, analysis results, documentation, software

Challenges in Scientific Data Management

- Standardization of file storage
- Data formats
- Metadata
- Linking of raw data, processed data, analysis results, documentation, software
- Retrievability and searchability

File formats, file names and folders

- Filename and folder structure conventions

  Example:
  `/Photos/Holidays/Europe/2019/img1.jpg` or
  `/Photos/2019/Holidays/Europe/img1.jpg` ?

- No central file storage

- Vendor-lock-in / Proprietary file formats / Missing APIs (application programming interfaces)

- Undocumented file formats

Metadata

Many file formats don't provide metadata storage.

Popular workarounds:

- Store metadata in filename:
  `datafile_networksimulation_20190207_a120_b17_CRPGM_debug.dat`
- Non-standardized text file formats

→ How to search metadata?

"Find all data files which were recorded by Person X."

“Missing Link”

- I have the plot `plot_network.pdf`, where is the plotting script and the raw data?
- I have the data file `datafile_networksimulation_20190207_a120_b17_CRPGM_debug.dat`, but haven’t I plotted the data before?

Important Requirements

- Search functionality

Important Requirements

- Search functionality
- Ability to store every data format, at any file size

Important Requirements

- Search functionality
- Ability to store every data format, at any file size
- Possibility to store, link and retrieve raw data, processed data, analysis results and documentation

Important Requirements

- Search functionality
- Ability to store every data format, at any file size
- Possibility to store, link and retrieve raw data, processed data, analysis results and documentation
- Support all kinds of data analysis software, from simple scripts to high-level software

Important Requirements

- Search functionality
- Ability to store every data format, at any file size
- Possibility to store, link and retrieve raw data, processed data, analysis results and documentation
- Support all kinds of data analysis software, from simple scripts to high-level software
- Minimally invasive workflow

Important Requirements

- Search functionality
- Ability to store every data format, at any file size
- Possibility to store, link and retrieve raw data, processed data, analysis results and documentation
- Support all kinds of data analysis software, from simple scripts to high-level software
- Minimally invasive workflow
- Scientific environments change often: Need for flexible data model

CaosDB

Research Data Management during Data Analysis

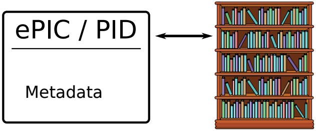
Data Aquisition

Electronic Lab Notebooks

z.B. RSpace, IDBS E-WorkBook, Biovia

Data Publication

Data Repositories



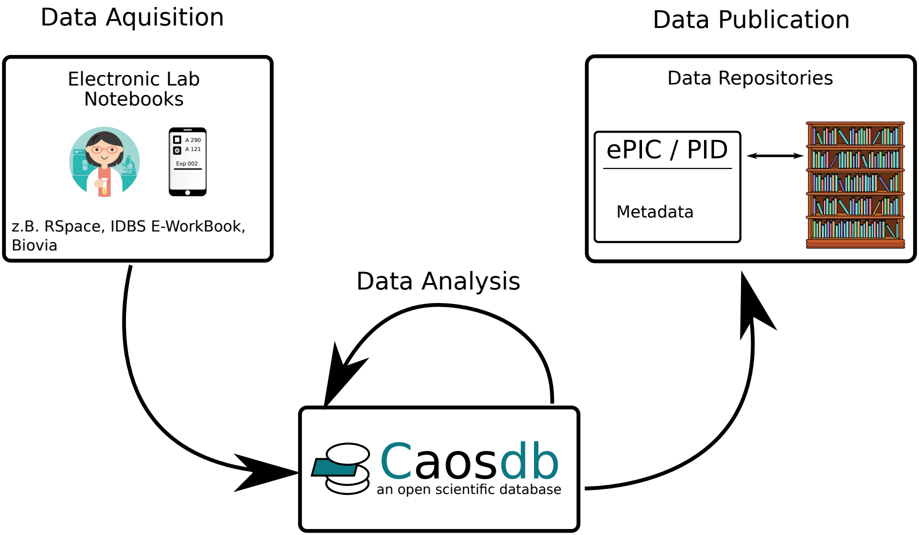
ePIC / PID

Metadata

Scientist picture: `201705 Scientist bench F.svg` from commons.wikimedia.org/wiki/Category:Life_science_images_from_DBCLS, CC-BY 4.0

Bookshelf: https://openclipart.org/detail/289378/bookshelf-with-blue-books

Research Data Management during Data Analysis



Data Aquisition

Electronic Lab Notebooks

z.B. RSpace, IDBS E-WorkBook, Biovia

Data Publication

Data Repositories

ePIC / PID

Metadata

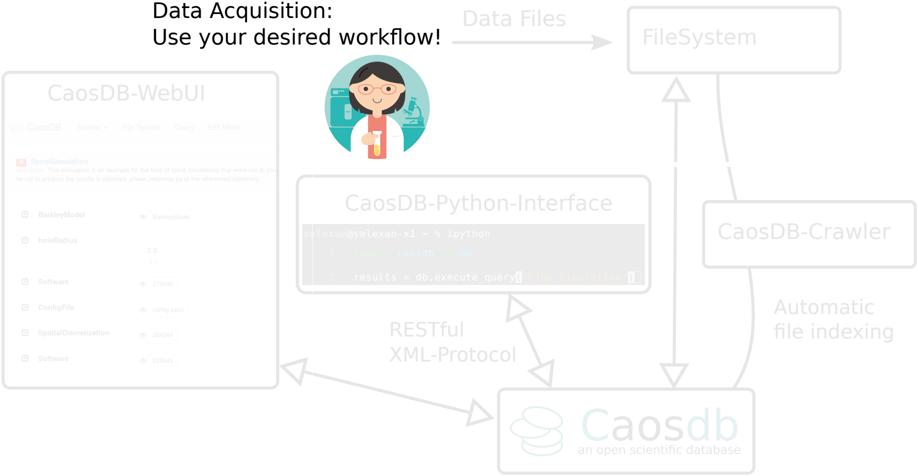
Data Analysis

Caosdb
an open scientific database

Scientist picture: 201705 Scientist bench F.svg from commons.wikimedia.org/wiki/Category:Life_science_images_from_DBCLS, CC-BY 4.0

Bookshelf: https://openclipart.org/detail/289378/bookshelf-with-blue-books

CaosDB Overview



Data Acquisition:
Use your desired workflow!

CaosDB Entities ▾ File System Query Edit Mode

R SpiralSimulation
description: This simulation is an example for the kind of spiral simulations that were run to pro
be run to produce the results is calculate_phase_response.py of the referenced repository.

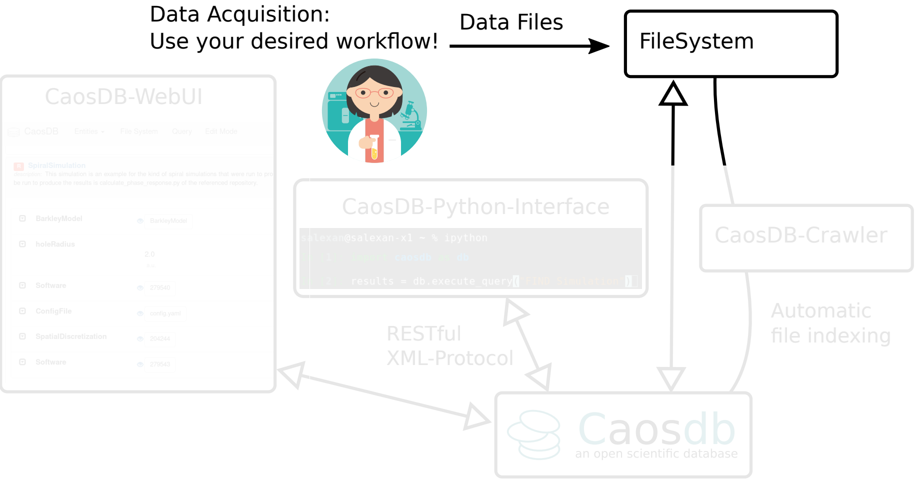
BarkleyModel	BarkleyModel
holeRadius	2.0 a.u.
Software	279540
ConfigFile	config.yaml
SpatialDiscretization	204244
Software	279543

```
salexan@salexan-x1 ~ % ipython
In [1]: import caosdb as db

In [2]: results = db.execute_query("FIND Simulation")
```

OpenSource! https://gitlab.gwdg.de/bmp-caosdb

CaosDB Overview



Data Acquisition:
Use your desired workflow!

R SpiralSimulation

description: This simulation is an example for the kind of spiral simulations that were run to pro
be run to produce the results is calculate_phase_response.py of the referenced repository.

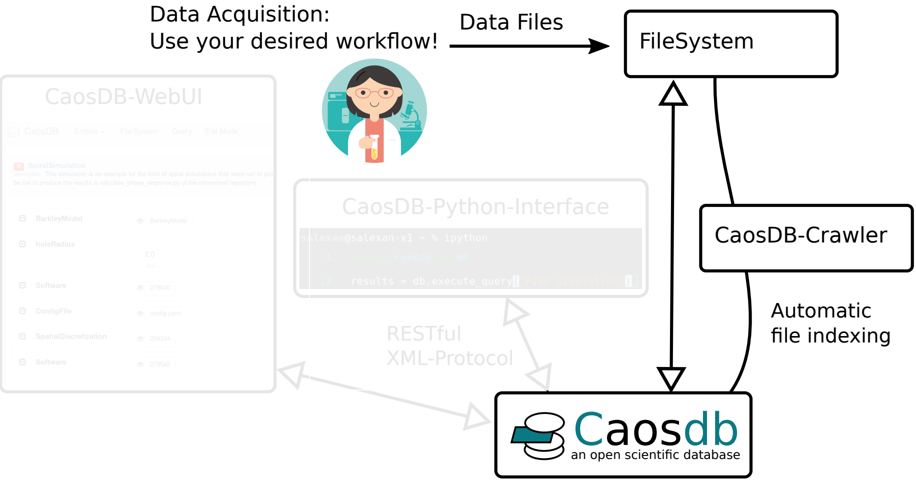
BarkleyModel	BarkleyModel
holeRadius	2.0 a.u.
Software	279540
ConfigFile	config.yaml
SpatialDiscretization	204244
Software	279543

```
salexan@salexan-x1 ~ % ipython
In [1]: import caosdb as db

In [2]: results = db.execute_query("FIND Simulation")
```

OpenSource! https://gitlab.gwdg.de/bmp-caosdb

CaosDB Overview



Data Acquisition:
Use your desired workflow!

Data Files

FileSystem

CaosDB-Crawler

Automatic
file indexing

CaosDB-WebUI

SpiralSimulation

description: This simulation is an example for the kind of spiral simulations that were run to pro
be run to produce the results is calculate_phase_response.py of the referenced repository.

BarkleyModel	BarkleyModel
holeRadius	2.0 a.u.
Software	279540
ConfigFile	config.yaml
SpatialDiscretization	204244
Software	279543

CaosDB-Python-Interface

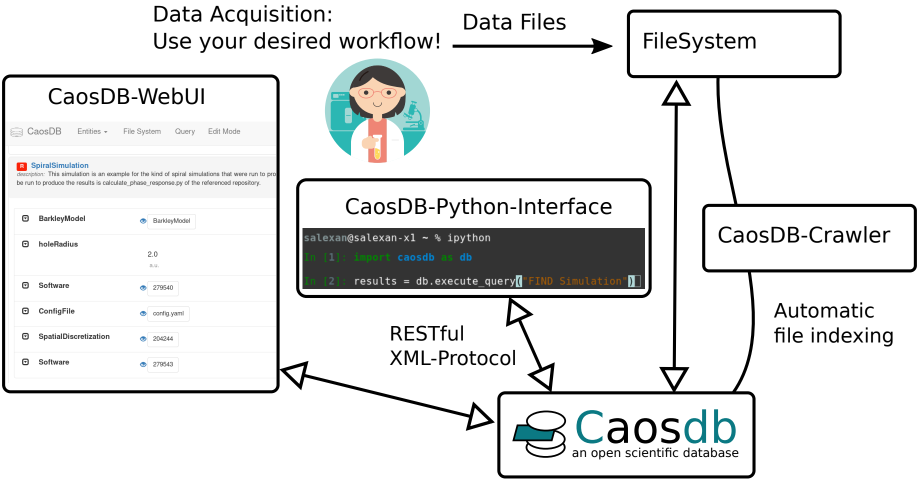
```
salexan@salexan-x1 ~ % ipython
In [1]: import caosdb as db

In [2]: results = db.execute_query("FIND Simulation")
```

RESTful
XML-Protocol

OpenSource! https://gitlab.gwdg.de/bmp-caosdb

CaosDB Overview



R SpiralSimulation
description: This simulation is an example for the kind of spiral simulations that were run to pro
be run to produce the results is calculate_phase_response.py of the referenced repository.

BarkleyModel	BarkleyModel
holeRadius	2.0 a.u.
Software	279540
ConfigFile	config.yaml
SpatialDiscretization	204244
Software	279543

```
salexan@salexan-x1 ~ % ipython
In [1]: import caosdb as db
In [2]: results = db.execute_query("FIND Simulation")
```

OpenSource! https://gitlab.gwdg.de/bmp-caosdb

State and Future of the Project

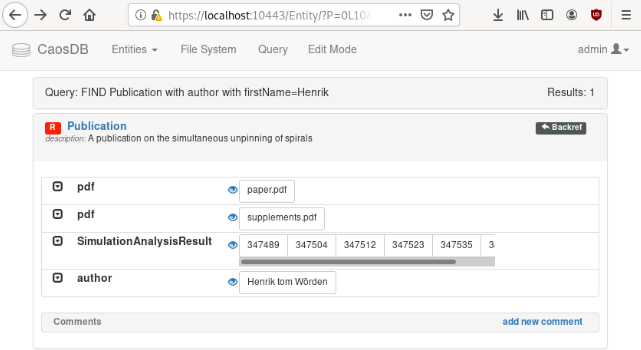
- CaosDB is a scientific project at the Research Group Biomedical Physics
- It is developed since ≈9 years and running stable since ≈2016
- CaosDB has been released as OpenSource software in 2018
- CaosDB is currently tested in several other workgroups in and outside of Göttingen
- Since May 2019: Commercial support and development by IndiScale GmbH

Demo: Tracing a publication back to its data

→ Query: FIND Publication with author with firstname = Henrik

Demo: Tracing a publication back to its data

→ Query: FIND Publication with author with firstname = Henrik



Query: FIND Publication with author with firstName=Henrik   Results: 1

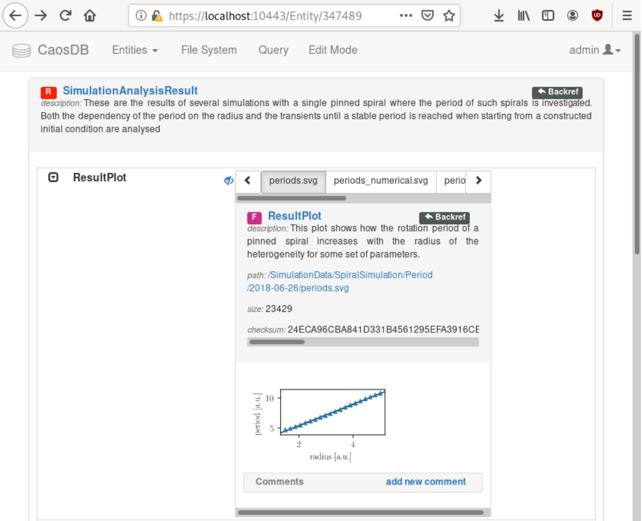
R **Publication**

description: A publication on the simultaneous unpinning of spirals

pdf	paper.pdf
pdf	supplements.pdf
SimulationAnalysisResult	347489 · 347504 · 347512 · 347523 · 347535 · 3
author	Henrik tom Wörden

See data policy statement here

Demo: Tracing a publication back to its data



SimulationAnalysisResult

description: These are the results of several simulations with a single pinned spiral where the period of such spirals is investigated. Both the dependency of the period on the radius and the transients until a stable period is reached when starting from a constructed initial condition are analysed

ResultPlot

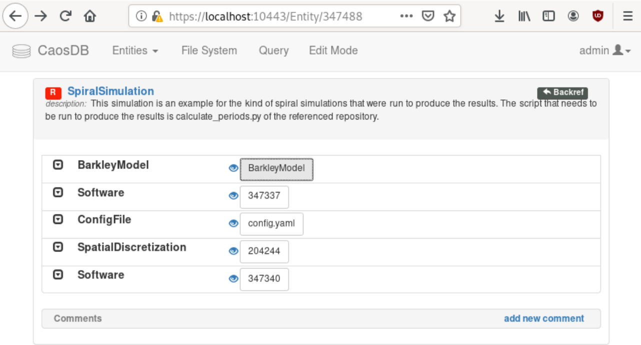
description: This plot shows how the rotation period of a pinned spiral increases with the radius of the heterogeneity for some set of parameters.

path: /SimulationData/SpiralSimulation/Period/2018-06-26/periods.svg

size: 23429

checksum: 24ECA96CBA841D331B4561295EFA3916CE

Demo: Tracing a publication back to its data



https://localhost:10443/Entity/347488

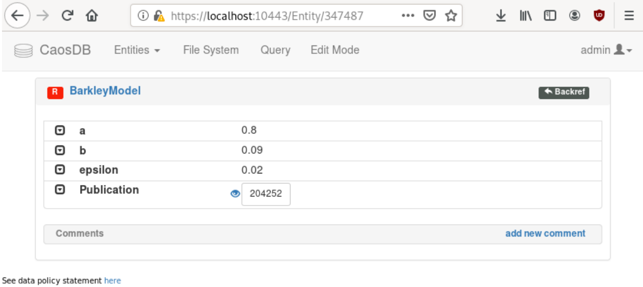
R **SpiralSimulation**

description: This simulation is an example for the kind of spiral simulations that were run to produce the results. The script that needs to be run to produce the results is calculate_periods.py of the referenced repository.

BarkleyModel	BarkleyModel
Software	347337
ConfigFile	config.yaml
SpatialDiscretization	204244
Software	347340

See data policy statement here

Demo: Tracing a publication back to its data



BarkleyModel

a	0.8
b	0.09
epsilon	0.02
Publication	204252

Comments add new comment

See data policy statement here

More Information

Caosdb
an open scientific database

- OpenSource Project: https://gitlab.com/caosdb
- Paper in Data https://doi.org/10.3390/data4020083
- Demo Instance (hosted by IndiScale) https://demo.indiscale.com (Beta test!)

Thank You!