

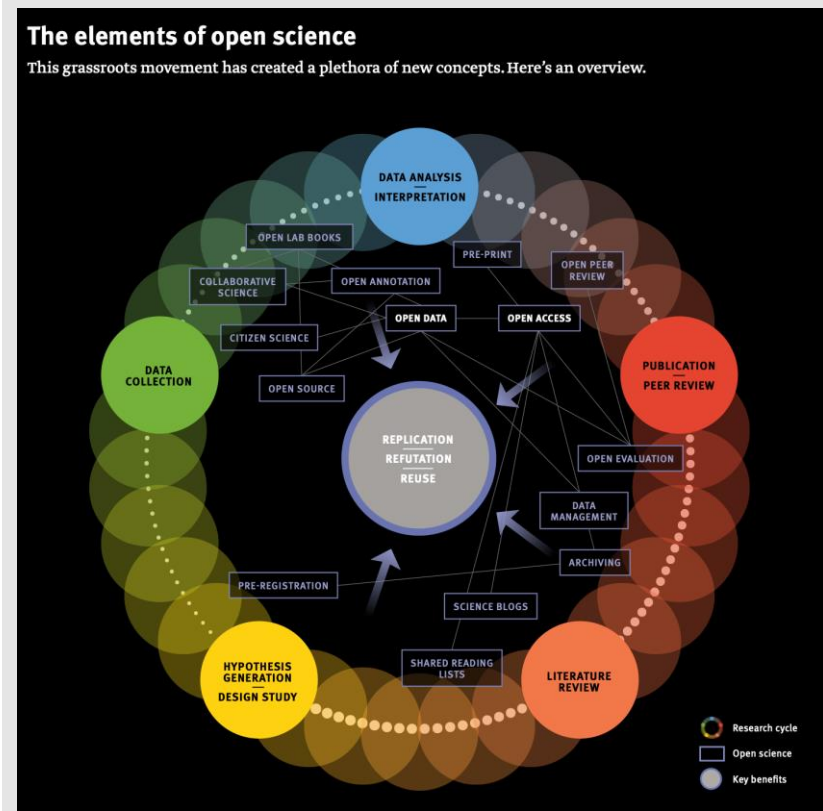
PAUL SCHERRER INSTITUT



Thorsten Bartels-Rausch :: Scientist :: Paul Scherrer Institut

Benefit of open research data in atmospheric science

ORD Meeting



«There is, however, a real problem with open lab notebooks: they can eat up too much time.»

SNF Horizonte 2016

Current Status :: Publication & data

EGU European Geosciences Union Atmospheric Chemistry and Physics

Home Search

OPEN ACCESS INTERACTIVE PUBLIC PEER REVIEW ARTICLE LEVEL METRICS

Executive editors: Ken Carslaw & Barbara Ervens
eISSN: ACP 1680-7324, ACPD 1680-7375

Atmospheric Chemistry and Physics (ACP) is a not-for-profit international scientific journal dedicated to the publication and public discussion of studies investigating Earth's atmosphere and the underlying chemical and physical processes. ACP publishes studies with important implications for our understanding of the state and behaviour of the atmosphere and climate, including the troposphere, stratosphere, and mesosphere.

Topics include gases, aerosols, clouds, precipitation, dynamics, radiation and their role in the Earth's climate system (including the biosphere, hydrosphere, and cryosphere). Research activities include laboratory studies, field measurements, remote sensing, modelling and data analysis, and machine learning (for details see journal subject areas).

Transparent peer review for 20 years: for 20 years, Atmospheric Chemistry and Physics has been a pioneer in transparent peer review. Submitted preprints, reviews, and author replies are posted and permanently archived on the journal website. This unique approach ensures the highest levels of scientific transparency and integrity, as well as [fair peer review](#) for authors.

- ✓ FAIR and open
- ✓ DOI links data and publications
- ✓ Extended published data
- ⚠ Connection to raw data and lab notes not clear

Map showing research locations in the Alps region, including Bern, Ticino, Lugano, and Trento. The map is overlaid with a yellow box containing the list of research topics.

Data set on interfacial supercooling and the precipitation of hydrohalite in frozen NaCl solutions b...

Laboratory experiments are presented on the phase change at the surface of sodium chloride – water mixtures at temperatures between 259 K and 240 K. High selectivity to the upper few nanometres of the frozen solution – air interface is achieved by using electron yield near-edge X...

An ice oxygen K-edge NEXAFS spectroscopy data set on gas-phase processing

Data are compiled that have been used to demonstrate the impact of high water partial pressure on X-ray absorption spectra of ice.

Data set on bromide oxidation by ozone in snow during metamorphism from laboratory study

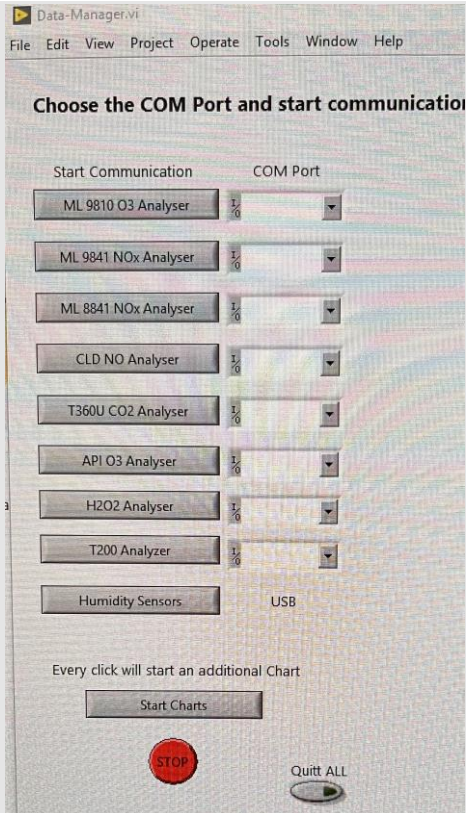
Earth's snow cover is very dynamic on diurnal time scales. The changes to the snow structure during this metamorphism have wide ranging impacts such as on avalanche formation and on the capacity of surface snow to exchange trace gases with the atmosphere. Here, we investigate the...

ATMOSPHERE CHEMISTRY ICE
LABORATORY STUDY
SNOW-ATMOSPHERE INTERACTIONS

ANTARCTICA ARCTIC ICE
SEA SALT AEROSOL SNOW

ALPINE ANTARCTICA ARCTIC
BRINE CHEMISTRY
HETEROGENEOUS INTERFACE
LABORATORY STUDY METAMORPHISM

Current Status :: Data



> H2	
> HU	
> IC/	
> Ko	
> Lo	
> MK	
> ML	
▼ ML	
> 2010	19.08.2016, 09:27
> 2011	19.08.2016, 09:27
▼ 2012	Heute, 08:42
> 05_Mai	19.08.2016, 09:27
> 06_Juni	19.08.2016, 09:27
> 08_August	19.08.2016, 09:27
▼ 09_September	19.08.2016, 09:27
2012-09-17_13.35.18_PC5010_ML9810_O3.dat	17.09.2012, 13:35
2012-09-17_13.36.02_PC5010_ML9810_O3.dat	17.09.2012, 15:01
2012-09-17_15.01.24_PC5010_ML9810_O3.dat	17.09.2012, 15:01
2012-09-18_15.23.55_PC5010_ML9810_O3.dat	18.09.2012, 15:24
> 11_November	19.08.2016, 09:27
> 2013	19.08.2016, 09:27
> 2014	19.08.2016, 09:27
> 2015	19.08.2016, 09:27
> 2017	13.12.2017, 09:43
> 2018	24.06.2019, 09:08
> 2020	01.12.2020, 10:44
data_2411 > UV_VIS__Varian_Cary50	



- Automated data storage
- Automated backup
- Date based naming convention



- Meta data & data not linked
- Note books not linked
- Only LabView data

Current Status :: Analysis

Präsentationen			
fits.pptx	30.11.2016, 08:11	390 KB	PowerP...n (.pptx)
Tabellen			
KLinC	26.04.2022, 14:04	12 KB	Microso...k (.xlsx)
experiment conditions	26.04.2022, 14:04	11 KB	Microso...k (.xlsx)
TUBE_CKD.XLS	26.04.2022, 14:04	224 KB	Microso...ok (.xls)
KLinC calc.xlsx	08.02.2017, 09:38	12 KB	Microso...k (.xlsx)
PDF-Dokumente			
Adsorption of H2O2 to ice -- KLinC	26.04.2022, 14:04	253 KB	PDF-Dokument
Adsorption of H2O2 to ice -- Ads vs. Des	26.04.2022, 14:04	234 KB	PDF-Dokument
Hong_Paper.pdf	11.08.2020, 16:49	6 KB	PDF-Dokument
SOFLA_101_118102416500.pdf	24.10.2018, 16:59	720 KB	PDF-Dokument
SOFLA_101_118102416510.pdf	24.10.2018, 16:58	750 KB	PDF-Dokument
SOFLA_101_118102416511.pdf	24.10.2018, 16:58	643 KB	PDF-Dokument
SOFLA_101_118102416512.pdf	24.10.2018, 16:58	540 KB	PDF-Dokument
Adsorption of H2O2 to ice -- KLinC Kopie.pdf	07.11.2016, 16:51	253 KB	PDF-Dokument
Bilder			
Fig2_Igor_V3_20160404_bkgcorr_Graph4.png	26.04.2022, 14:04	119 KB	PNG-Bild
Fig4_Igor_V3_bt_fits_Graph2_1.png	26.04.2022, 14:04	495 KB	PNG-Bild
Fig5_Igor_V3_int uptake recovery_BKGCORR_Graph0_4.png	26.04.2022, 14:04	92 KB	PNG-Bild
Hong_Paper.png	11.08.2020, 16:49	66 KB	PNG-Bild
Entwickler			
fit_background_H2O2.m	26.04.2022, 14:04	4 KB	Objecti...uellcode
FlowTube_Angela_Thomas.m	26.04.2022, 14:04	5 KB	Objecti...uellcode
FT_H2O2_20210609.m	26.04.2022, 14:04	8 KB	Objecti...uellcode
SurfaceCoverageBartelsRausch2013.m	26.04.2022, 14:04	1 KB	Objecti...uellcode
Andere			
matlab_Angela_Thomas_Flowtubes.mat	26.04.2022, 14:04	128 Byte	MATLAB Data
V3_normalised uptake_BKGCORR.xpx	23.10.2018, 18:01	698 KB	Igor Pro...eriment
V3_int uptake recovery_BKGCORR.xpx	15.03.2018, 13:55	2.7 MB	Igor Pro...eriment
V3_bt fits.xpx	15.03.2018, 13:51	514 KB	Igor Pro...eriment
V3_20160404_bkgcorr.xpx	15.03.2018, 13:47	4.2 MB	Igor Pro...eriment

- Individual notes, internal reports, presentations & manuscripts



- Variety of file types
- No naming convention
- No version control
- No link to data or publication
- No structure

2022_AFPXS_BartelsRausch_Napp.pptx	06.12.2022, 10:13	06.12.2022, 08:56
Bartels-Rausch_SGM_2017_Davos.pptx	26.04.2022, 13:39	17.11.2017, 12:40
Bartels_GenLab_Sapporo_2018.pptx	26.04.2022, 13:38	23.09.2018, 09:47
GruSi_FAdosing.pptx	26.04.2022, 13:38	09.01.2017, 09:41
20151015_Bartels_NappScience_TestExp.pptx	26.04.2022, 13:37	30.12.2015, 21:16
Poster_hoch_GAC_2018.pptx	26.04.2022, 13:37	17.09.2018, 13:47
Bartels_GruMe_201711_DP_Phoenix.pptx	26.04.2022, 13:37	18.11.2017, 11:02
Bartels_Poster_hoch_PCL_Napp.pptx	26.04.2022, 13:36	03.01.2018, 11:31
Bartels-Rausch_PCI_2018_Zürich.pptx	26.04.2022, 13:36	11.01.2018, 11:19
Bartels-Rausch_UoT_2018.pptx	26.04.2022, 13:36	10.07.2018, 13:55
GroupMeetingTomOct	26.04.2022, 13:35	30.12.2015, 21:16
DepthProfilesAtSIM.pptx	26.04.2022, 13:34	16.09.2016, 12:02
Thorsten_Highlight_HCL_Ice.pptx	26.04.2022, 13:33	23.11.2017, 08:36
20151016_Bartels_SAOG_NaCl.pptx	26.04.2022, 13:32	22.01.2016, 12:52
LCH_Seminar	26.04.2022, 13:31	30.12.2015, 21:16
Bartels-Rausch_ENE-LUC-Highlight_HCL_Ice.pptx	26.04.2022, 13:31	21.09.2017, 16:15
AGU_2016.pptx	26.04.2022, 13:31	10.12.2016, 09:55
Poster_hoch_GruSi_Napp.pptx	26.04.2022, 13:31	18.05.2017, 13:26

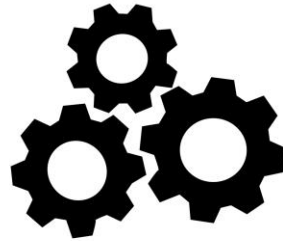
F
Findable



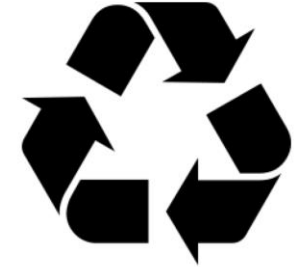
A
Accessible



I
Interoperable



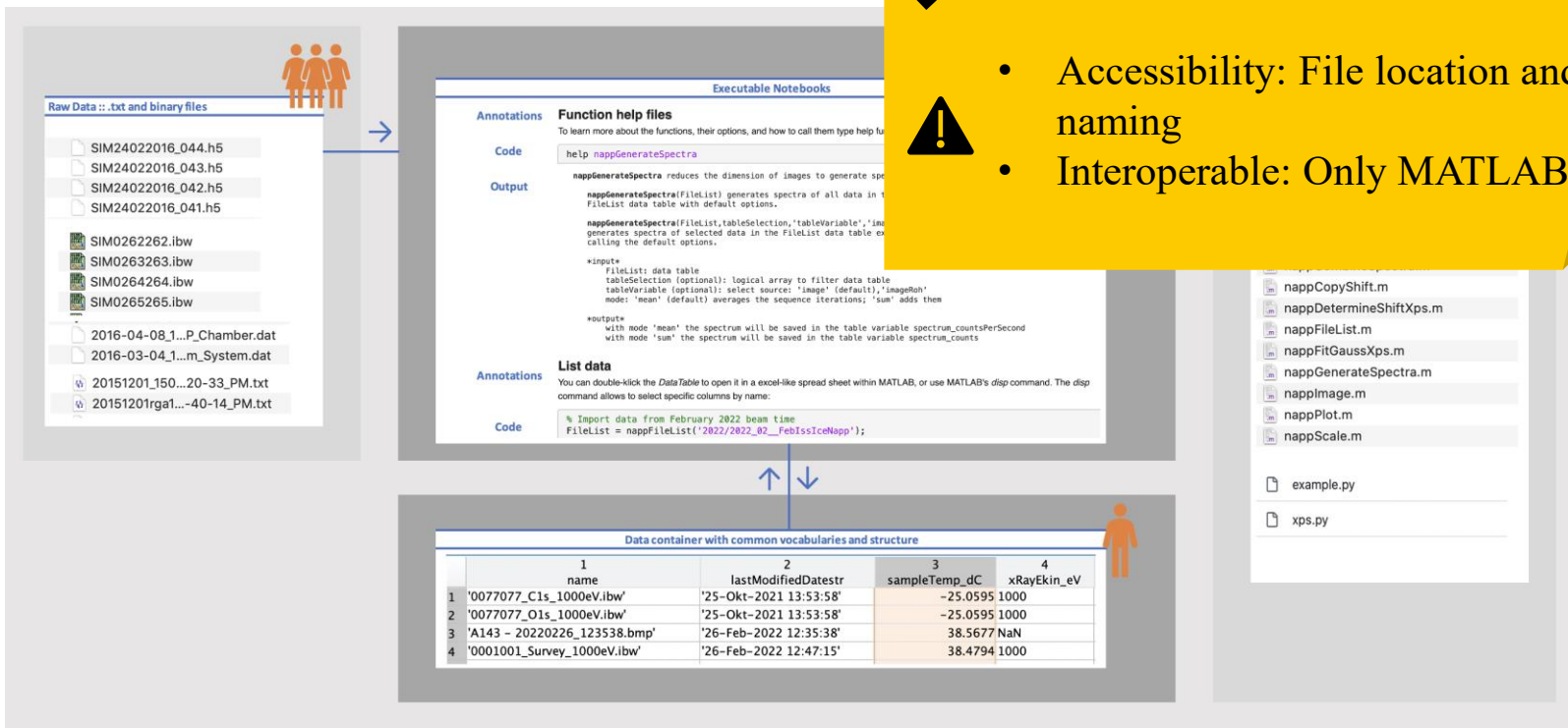
R
Reusable



- Link data – analysis – publication
- Make data & meta data & notes findable
- The human factor: Automate data linking and storage location.

- Interoperable <-> naming convention
- Findable and Interoperable: Controlled vocabulary
- Accessible: Import procedures

Connecting meta data & data & notes



- ✓ Findable data & meta data & notes
- ⚠ Accessibility: File location and naming
- Interoperable: Only MATLAB

Lab notebook: key features

- 1. Links to samples and protocols.** In the description of experimental procedures it is possible to establish and annotate links to samples and protocols stored in the inventory.

Sample: Detection of LexA-ER-B42 induction by flow cytometry
 /DIANA_OTTOZ/INDUCIBLE_TRANSCRIPTION_FACTOR/INDUCTION_OF_TF/FC_LEXA-ER-B42

General

Name: Detection of LexA-ER-B42 induction by flow cytometry
Owner: Diana Ottos

Experimental goals: Analyze the induction of LexA-ER-B42 in a concentration series of beta-estradiol using a fluorescence readout.
Experimental results: The LexA-ER-B42 induction can be measured by using a target gene encoding a fluorescence protein, LexA-ER-B42 induction is different at different concentrations of inducer.

Parents

Identifier	Name	Annotation :: Comments	Space	Modification Date
/MATERIALS/FRM11	liquid B media	SDC at 20°C	MATERIALS	2015-07-06 16:15:22
/MATERIALS/FR060634	1000X Cycloheximide	diluted 1/1000, treated for 1/2 hour	MATERIALS	2015-07-06 16:13:48
/MATERIALS/FR060637	beta-estradiol	1:2 concentration series with 3000 nM maximum for 24 hours	MATERIALS	2015-07-06 16:13:48
/MATERIALS/FRY416	LexA-ER-B42 + target		MATERIALS	2015-07-17 12:07:15
/METHODS/FRPROT26	Flow cytometry to analyze fluorescence levels in single tubes (Basic)	channels used: 488 for FSC (A, W) and SSC (A, W), 488 (B0030 A), 661 (B1020 A)	METHODS	2015-04-11 11:45:20

1 - 5 of 5 items | 10 | Per Page | Page 1 | of 1

Children

Readout details

Experimental readout: flow cytometry
Machine: LSRI Fortessa

Comments

Publication: Ottos et al., Nucleic Acids Research, 2014

FRY416 24 h induction

RAW_DATA : 2015041113151319-234
 ANALYSIS_SCRIPTS : 20150411131856330-235
 ANALYZED_DATA : 20150411131831277-236
 ELN_PREVIEW : 20160813115814516-324

Files Uploader

Select files to upload

Create Dataset
 Auto upload on drop

- 2. Links to other experimental steps**

Jupyter notebook examples

Cancer genomics analysis

ucsd-ccb / jupyter-genomics

Watch 14 Star 43 Fork 18

Code Issues 0 Pull requests 0 Projects 0 Wiki Security Insights

Branch: master jupyter-genomics / notebooks / networkAnalysis / TCGA_clustering_OV / TCGA_clustering_OV.ipynb Find file Copy path

brinrosenthal added network analysis notebooks 8a5d87e on 10 Jun 2016

1 contributor

2.62 MB Download History

Clustering methods applied to TCGA Ovarian Cancer Coexpression Matrix

Author: Brin Rosenthal (sbroenthal@ucsd.edu)

April 15, 2016

- We will provide a first look into which clustering methods work best on TCGA coexpression matrices, starting with Ovarian Cancer.
- We will visualize these clustering results.
- Clustering methods tested include:
 - Modularity maximization (Louvain)
 - Affinity propagation
 - DBSCAN
 - Hierarchical clustering
- The Ovarian Cancer TCGA Co-expression matrix may be found here https://ucsd-ccb-data-analysis.s3.amazonaws.com/Brin/ccbb_jupyter_genomics/network_analysis/brin_OV_clustering_TCGA/OV.tsv

Plot the clusters in network form

- Spring embedded layout for node positions (more strongly connected nodes are positioned closer together)
- Node colors encode cluster membership found from the Louvain modularity maximization algorithm

```
In [1]: # plot the network
import matplotlib.colorbar as cb
import seaborn as sns

vmin=None
vmax=None

cmap = 'Paired'

pos = nx.spring_layout(Gtemp,k=.03)

fig,ax=plt.subplots(figsize=(50,40))

# draw small community nodes as white
partition = pd.Series(partition)
par_VC = partition.value_counts()
groupL5 = list(par_VC[par_VC<5].index)
groupG5 = list(par_VC[par_VC>=5].index)

# select out nodes in small communities
nodes_w = []
[nodes_w.extend(list(partition[partition==i].index)) for i in groupL5]

# now select large community nodes
nodes_c = []
[nodes_c.extend(list(partition[partition==i].index)) for i in groupG5]

GL5 = nx.subgraph(Gtemp,nodes_w)
GG5 = nx.subgraph(Gtemp,nodes_c)

# rename partitions for plotting
group_map = dict(zip(groupL5,range(1,5)))
par_rename = []
[par_rename.append(group_map[i]) for i in groupL5]

cols = par_rename
cols = pd.Series(cols,index=nodes_w)
#cols[nodes_c]=partition
#cols = list(cols)

nx.draw_networkx_nodes(Gtemp,nodes_w,pos,ax=ax,edgecolors='black',node_color=cols,edge_cmap=cmap,edge_vmin=vmin,edge_vmax=vmax)

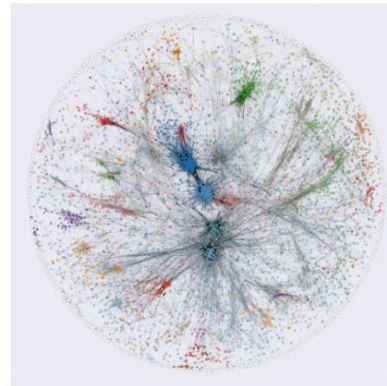
nodes_col = nx.draw_networkx_nodes(Gtemp,nodes_c,pos,ax=ax,edgecolors='black',node_color=cols,edge_cmap=cmap,edge_vmin=vmin,edge_vmax=vmax)

nx.draw_networkx_edges(Gtemp,pos,ax=ax,edge_cmap=cmap,edge_vmin=vmin,edge_vmax=vmax)

plt.grid('off')
```

Run the above code block to generate network figure

- Static image included here to conserve space




ORD Project :: IVDAY

EVOLUTION of a research project

Versatile data input and collaborative analysis



- PSI: openBIS data import
 - SLS: vocabulary and file format
 - PSI: Procedures in openBIS
- 
- **Find**
 - Data
 - r
 - Integrate GUI for code
 - **Interoperable:**
 - Define Hand-over points
 - Define data format and structure for those (netCDF, h5)
 - **Accessible:**
 - Establish a process to effectively explore and visualize data at any stage of the FAIR research cycle.



- 1 scientific programmer
- 3 research groups
- 6 science cases
- 30 researchers
- Wide range of instrument types (shared among groups)
- International Outreach
- Customize openBIS with report functionality

Thanks

