

Ich hab es hier irgendwo

Grundlagen der Arbeit mit Daten

Felix Rau & Patrick Helling
(Universität zu Köln)



Data Center for the Humanities
Kölner Datenzentrum
für die Geisteswissenschaften

Worum geht's?

Kontrolle beim Umgang mit Daten.

Fahrplan

Zeit		Inhalt
14:00 Uhr	5 min	Intro
14:05 Uhr	30 min	Daten & Co.
14:35 Uhr	10 min	FDM Prinzipien
14:50 Uhr	45 min	Workflows
15:20 Uhr	20 min	Diskussion 
15:50 Uhr	10 min	Abschluss

Wer sind wir


Felix Rau



Patrick Helling

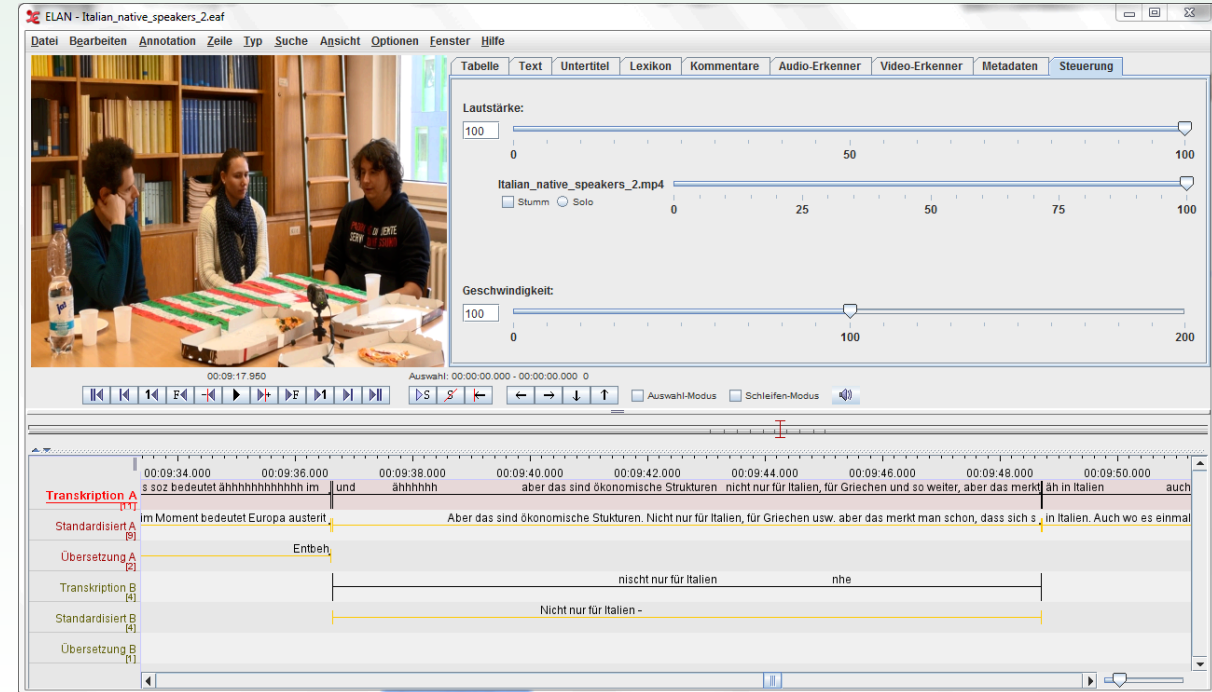


Daten & Co.

Dauer	Inhalt
5 min	Daten
5 min	Metadaten
5 min	Paradaten
5 min	Anderer Output
10 min	Diskussion 

Daten

In der Linguistik:
Repräsentation eines
gesprochenen,
gebärdeten oder
geschriebenen
Sprachevents.



Datentypen

Literatur 📖, Bilder 🖼️, Audio 🎵, Fotos 📷,
Datenbanken 💿, Videos 🎥, Objekte 🏺, Eye tracking 👁️,
Tabellen 📊, Manuskripte 📝, Software 💾, fmri-
Daten 🧠, ...

Daten

Leonelli 2016 defines data as

“ any product of research activities, ranging from artifacts such as photographs to symbols such as letters or numbers, that is collected, stored, and disseminated in order to be used as evidence for knowledge claims. ”

Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study*. Chicago ; London: The University of Chicago Press.

Datenmodelle

Sprachliche Interaktionssituation → Videoaufnahme
→ Transkript → Übersetzung → Morphem-
Glossierung → Intonationsannotation
30 Aufnahmen 🖐️ ⇒ numerischer Datensatz
(Morphem-Intonations-Korrelation)

Daten und Datenmodelle

- Welche Daten habe ich?
- Welche Derivate (Transformationen/Annotationen) habe ich?
- Wie sind diese verschiedenen Datenmodelle verknüpft?

Metadaten

“ Metadaten sind strukturierte Informationen über Forschungsdaten, die das Sammeln, Speichern, die Analyse und Verbreitung der Daten ermöglichen und unterstützen. ”

Metadaten (Beispiele)

Home

Insert

Draw

Page Layout

Formulas

Data

Review

View

Automate

Acrobat

AutoSave

Save

Undo

Redo

Find

Print

Print Range

Print Selection

Print Comments

Print Formulas

Print All

Print Range

Print Selection

Print Comments

Print Formulas

Print All

Search (Cmd + Ctrl + U)

Comments

Share

Paste

Format as Table

Cell Styles

Conditional Formatting

Format as Table

Cell Styles

Insert

Delete

Format

Sort & Filter

Find & Select

Add-ins

Analysis Data

Create PDF and share link

36

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

zenodo.org/records/14983644

RDM Demerit Badges

Rau, Felix · Möllenhaus, GJ Elisabeth^{1,2}

Published March 6, 2025 | Version 1.0.0

The RDM Demerit Badges were developed by Felix Rau and Elisabeth Möllenhaus at the Data Center for the Humanities at the University of Cologne. The Demerit Badges are intended to raise awareness for common mistakes in research data management (RDM) and to encourage good RDM practices as well as a positive culture around mistakes. The badges were designed to be used in workshops, lectures, and other teaching or outreach settings. The home of the RDM Demerit Badges is <https://dch.phil-fak.uni-koeln.de/rdm-demerit>.

All materials are licensed under the CC-BY 4.0 license.

See also:
Rau, F. & Möllenhaus, E. (2025). Fehlerkultur und die Vermittlung von Forschungsdatenmanagement. D44 2025 Under Construction (D442025). Bielefeld, Deutschland. <https://dx.doi.org/10.5281/zenodo.14943156>

File types and function

In this resource, we provide files that can be used to reproduce the stickers (image files) and files to adapt and modify the stickers (Adobe Illustrator files). The InfoSheets are intended to contextualize and explain the badges in an offline teaching setting. They are provided as PDF files for reproduction and as InDesign files for modification.

Badges

- Adobe Illustrator files for exporting JPEGs for printing ([badges_illustrator_print](#))
- Adobe Illustrator files for exporting web images ([badges_illustrator_web](#))
- JPEG files for printing stickers ([badges_png_print](#))
- PNG files for web pages ([badges_png_web](#))

InfoSheets

- InDesign files for exporting PDFs ([infosheets_indesign](#))
- PDF files for printing ([infosheets_pdf](#))

19 VIEWS · 17 DOWNLOADS

Versions

Version	Date
Version 1.0.0	Mar 6, 2025

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.14983643. This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

External resources

Indexed in

[OpenAIRE](#)

Communities

Paradaten

“ The paradata of a data set or survey are data about the process by which the data were collected. [Wikipedia](#) ”

Paradaten (Beispiele)

- Ursprünglich aus der Umfrageforschung
- Heute auch in anderen Disziplinen
- Beispiele: Zeitstempel, keystroke-Daten, Sprechtempo, Befragungszeitpunkt, ...
- Können als eine Art Metadaten betrachtet werden

Anderer Output

- Code
- Paper
- Was habe ich und wie ist es mit den Daten verknüpft?

Wie sieht es bei euch aus? ✨

Wie sieht es bei euch aus?

- Daten
 - Rohdaten
 - Derivate/Datenmodelle
- Metadaten
- Paradata
- Anderer Output

FDM Prinzipien und Workflows

Dauer Inhalt

10 min FDM Prinzipien

10 min Workflows

10 min Datensuche 

20 min Diskussion

FDM Prinzipien

- Dokumentiere Strukturen
- Dokumentiere Entscheidungen
- Dokumentiere die Änderungen (z.B. via Code)
- Versioniere Daten
- Automatisiere Prozesse
- Wiederholbarkeit (Repeatability)
- Kontrolle über Daten und Prozess

FAIR

<https://www.go-fair.org/fair-principles/>

- Findable
- Accessible
- Interoperable
- Reusable

Workflows

- Acquiring (erheben/erwerben)
- Keeping (Daten halten)
- Handling (handhaben)
- Releasing (veröffentlichen/löschen/archivieren)

Acquiring (erheben/erwerben)

- Dokumentieren, wie Daten erhoben wurden
 - Paradata
- Datenqualität
- Datenformate
- Rechtliche Aspekte

Datensuche

1. VLO (<https://vlo.clarin.eu/>)
2. BASE (<https://www.base-search.net/>)
3. FDB (<https://www.forschungsdaten-bildung.de/>)
4. Google Dataset Search
(<https://datasetsearch.research.google.com/>)
5. SSH Open Marketplace
<https://marketplace.sshopencloud.eu/>
6. R3Data (Repositories)
<https://www.re3data.org/search>
7. DataCite Commons <https://commons.datacite.org/>

Keeping (Daten halten)

- Speicherung
- Backup
- Dateitypen
- Ordnerstruktur
- Dateibenennung

Speicherung und Backup

- Speicherung
 - Lokal / Cloud / Institutionell
- Backup
 - 3-2-1 Regel
 - Automatisch
 - Regelmäßige Tests

Rau, F. (2023). *DCH Project Data Storage Guidelines* (1.0.0). Zenodo. [doi:10.5281/zenodo.7760967](https://doi.org/10.5281/zenodo.7760967)

Dateitypen

- Offene, standardisierte Formate
- Arbeits- und Archivierungsformate

Ordnerstruktur

```
project/  
├── 01_admin/  
├── 02_data/  
│   ├── 01_raw/  
│   └── 02_processed/  
├── 03_code/  
├── 04_docs/  
├── 05_output/  
├── README.md  
└── LICENSE
```

Rau, F. (2023). *DCH Folder Structure Guidelines* (1.0.0). Zenodo. [doi:10.5281/zenodo.7452113](https://doi.org/10.5281/zenodo.7452113)

Dateibenennung

1. Vergib eindeutige, aussagekräftige, lesbare, aber prägnante Namen
2. Wähle Namen, die in allen Datei- und Betriebssystemen sicher sind
3. Strukturiere den Dateinamen und nutze Dateinamenerweiterungen
4. Erleichtere die alphabetische Sortierung
5. Dokumentiere das Benennungsmuster

Beispiel Dateibenennung

[ISO 639-3]_[speaker]_[date]_[number]_[data type].[file extension]

ger_LK_2023-05-31_001_raw-audio.wav

ger_LK_2023-05-31_001_raw-video.m4v

Rau, F. (2023). DCH File Naming Guidelines (1.0.0). Zenodo. [doi:10.5281/zenodo.7447485](https://doi.org/10.5281/zenodo.7447485)

Handling (handhaben)

- Versionierung
- Dokumentation der Transformationen und Analyse
- Alles in Code
- Automatisierung
- (Analyse)

Releasing (veröffentlichen/löschen/archivieren)

- Datenarchivierung
- Datenpublikation
- Datenlöschung

Datenarchivierung

- Metadaten
- Datenstandards
 - Nachhaltigkeit der Formate
- 10 Jahre Archivierung (Mindestanforderung)
- Langzeitarchivierung

Datenpublikation

- Repositoriumssuche
- Metadatenstandards

Repositoryumssuche

- OSF (<https://osf.io/>)
- Zenodo (<https://zenodo.org/>)
- Fachspezifische Repositorien
 - R3Data (<https://www.re3data.org/search>)
 - <https://www.konsortswd.de/angebote/forschende/alle-datenzentren/>
- Institutionelle Repositorien (Lokale Einrichtungen)

Eure Perspektive

- Wie organisiert ihr eure Daten?
- Welche Herausforderungen seht ihr?
- Welche Best Practices könnt ihr teilen?
- Gibt es offene Fragen oder Unsicherheiten?
- Wie könnte ein idealer Workflow für euch aussehen?
- Welche Tools oder Ressourcen nutzt ihr bereits?

Danke!

info-dch@uni-koeln.de



Data Center for the Humanities
Kölner Datenzentrum
für die Geisteswissenschaften

