

*TH32F: Publishing Research Data: Peer Review, Data Center
Accreditation, and Linking*

Peer REview for Publication & Accreditation of Research data in the Earth sciences (PREPARDE)

Jonathan Tedds (Leicester), Sarah Callaghan (BADC), Fiona Murphy
(Wiley), Matt Mayernik (NCAR)

jat26@le.ac.uk

#PREPARDE

Data Publishing Issues & Workflows

- Why?
- What?
- Where?
- How?

Why publish data?

- Provide academic credit to scientists who create and manage data
 - without diverting effort from their main work on ensuring data quality.
- Ensures that the dataset is uploaded to a trusted repository where it will be archived and curated properly.
- Peer-review process also reassures the funder that the published dataset is of good quality and that the experiment was carried out appropriately.
- Allows researchers outside the immediate field, to discover data and learn access conditions.
- Shows transparency in the scientific process, improving public accountability.

Geoscience Data Journal, Wiley Blackwell and the Royal Meteorological Society

- **supported by NERC** – in particular the **British Atmospheric Data Centre**
- Partnership formed between **Royal Meteorological Society** and academic publishers **Wiley Blackwell** to develop a mechanism for the formal publication of data in the Open Access *Geoscience Data Journal*
- **GDJ publishes** short data papers cross-linked to, and citing, datasets that have been deposited in approved data centres and awarded DOIs (or other permanent identifier).
- A **data article** describes a dataset, giving details of its collection, processing, software, file formats etc, without the requirement of novel analyses or ground breaking conclusions. It **allows the reader to understand the when, how and why data was collected and what the data-product is.**

JOURNAL TOOLS

- Get New Content Alerts
- Get RSS feed
- Save to My Profile
- Recommend to Your Librarian

JOURNAL MENU

Journal Home

FIND ISSUES

Current Issue
All Issues
Virtual Issues

FIND ARTICLES

Early View
Editors' Choice

FOR CONTRIBUTORS

Author Guidelines
Submit an Article

ABOUT THIS JOURNAL

Editorial Board
News
Overview
Permissions
Advertise
Contact

SPECIAL FEATURES

Virtual Issue - Evolutionary Applications to Climate Change

For Reviewers

Cover Gallery

Wiley Job Network

Institutional and Funder Payments

Wiley Open Access

Open Access License and Copyright

Article Publication Charges

Virtual Special Issue - Geoengineering

Virtual Issue - Editor's Choice papers

L F Richardson Award Prize Winners

Geoscience Data Journal
Open Access

Editor in Chief
Dr Rob Allan,
Met Office, UK

➤ An earth science
open access
data journal



Data Paper

On the South Atlantic Convergence Zone affecting southern Amazonia in austral summer

Fabien C. Lamaze^{1,*}, Dany Garant², Louis Bernatchez¹

Article first published online: 23 OCT 2012

DOI: 10.1002/asl.401

© 2012 The Authors
This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

This is an Open Access article under the terms of the **Creative Commons Attribution Non Commercial License** which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Total views since publication: 25 [view chart](#)

Additional Information ([Show All](#))

[How to Cite](#) | [Author Information](#) | [Publication History](#) | [Funding Information](#)

Abstract

Article

References

Supporting Information

Cited By

Get PDF (504K)

Keywords:

brook charr; gene expression; hybridization; introgression; quantitative PCR; stocking

Dataset

Jump to...

GBS (Global Broadcast Service), [doi:10.1029/2007RS003793](https://doi.org/10.1029/2007RS003793)

Abstract

Jump to...

Time series analysis of the average rainfall over a target area in southern Amazon Basin showed a spectral peak at 11 day period. An objective method for defining the South Atlantic Convergence Zone (SACZ) is used to identify 28 episodes affecting southern Amazon Basin during the 10 summers in the period 1999–2010. The 28-episode composite precipitation anomalies show significant positive values over the target area. The convergence of moisture over the target area in the SACZ composites is about 35% stronger than the climatological value. Copyright © 2012 Royal Meteorological Society

1. Introduction

Jump to...

One of the regional scale meteorological systems that affect the weather over a major part of the South American tropics is the South Atlantic Convergence Zone (SACZ). This system somewhat plays the same role for the South American monsoon (Vera *et al.*, 2006; Carvalho *et al.*, 2010) as does the monsoon trough for the South Asian summer monsoon (Keshavamurty and Awade, 1970). The cloud band associated with SACZ extends from the Amazon Basin (Amazonia) to the South Atlantic subtropics, over a stretch of 4000 km or more (Kodama, 1992; Satyamurty *et al.*, 1998) and affects many regions of Brazil with intense rainfall. Some SACZ events are especially intense over interior South American continent (Carvalho *et al.*, 2002, 2004; Muza and Carvalho, 2006), strongly affecting Amazonia.



SEARCH

In this issue
Advanced > Saved Searches >

ARTICLE TOOLS

- Get PDF (504K)
- Save to My Profile
- E-mail Link to this Article
- Export Citation for this Article
- Get Citation Alerts
- Request Permissions

Share |

Guidelines & Workflows

For Authors:

- Are there guidelines on how to structure a Data Paper?
- How can I submit a Data Paper to ***Geoscience Data Journal***?
- What is a **Digital Object Identifier (DOI)** and why should datasets have them?
- What alternative permanent identifiers are allowable?

For Data Centers:

- How do data centers fit in to the **GDJ** ecosystem?
- What would my data center have to do to be approved by **GDJ**?
- Is the repository approval a lengthy process?
- Is there any support or guidance available from the journal? Whom do I contact to get my repository approved?
- Once approved what then?
- What are the benefits of encouraging researchers to submit to GDJ?

PREPARDE:

Peer Review for Data Publication

- Technical
 - author guidelines for GDJ
 - NERC Data Value Checklist
 - implicit peer review of repository?
- Scientific
 - pre-publication?
 - post-publication? E.g. F1000R
 - guidelines on uncertainty e.g. IPCC
 - discipline specific?
 - EU Inspire spatial formatting
- Societal
 - contribution to human knowledge
 - reliability

Peer review process for GDJ

- **data description document (data paper)**
 - describes the experiment in a reader-friendly way
 - may contain quick-look plots of the data
 - should describe how the data were created
 - so that the quality of the scientific method can be examined
 - provide information on the importance, uniqueness and applicability to other purposes of the data.
- **metadata**
 - clearly identify and describe the data.
- **The data themselves**
 - Usability
 - accessibility through the repository

Peer review process for GDJ

Review 1 – Data description document

- Is the method used to create the data of a high scientific standard?
- Is enough information provided (in metadata also) to enable the data to be re-used or the experiment to be repeated?
- Does the document provide a comprehensive description of all the data that is there?
- Does the data make an important and unique contribution to the geosciences?
- What range of applications to geosciences does it have?
- Are all contributors and existing work acknowledged?
- Does the Data Paper contain sufficient citation information of the dataset, eg dataset DOI, name of data centre etc. (see [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)2049-6060/homepage/ForAuthors.html#reference_style](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2049-6060/homepage/ForAuthors.html#reference_style) for necessary dataset citation details)

Peer review of data

- Survey of existing literature and journal data publication policies on-going
 - Discipline specific & general
- Workshop in UK early 2013
 - community input into suggested data review policy

What's out there already?

- IPCC guidelines on uncertainty
- NERC – HINDAWI
- US Geological Survey
- EU Inspire
- RE3data.org
- ESSD
- Societal implications
 - Contribution to human knowledge vs reliability
- Pensoft
- Life sciences
 - BMC medical
 - F1000Research
 - JOAD
 - ELife

Citation best practises

- International CODATA Task Group on Data Citation Standards and Practices
<http://www.codata.org/taskgroups/TGdatacitation/index.html>
 - White Paper on Current Practices in Data Citation (planned publication Dec 2012)
- SCOR/IODE/MBL WHOI Library Data Publication Working Group
- NERC data citation and publication project
- DataCite

Emerging Principles for Data Citation (1)

Derived from:

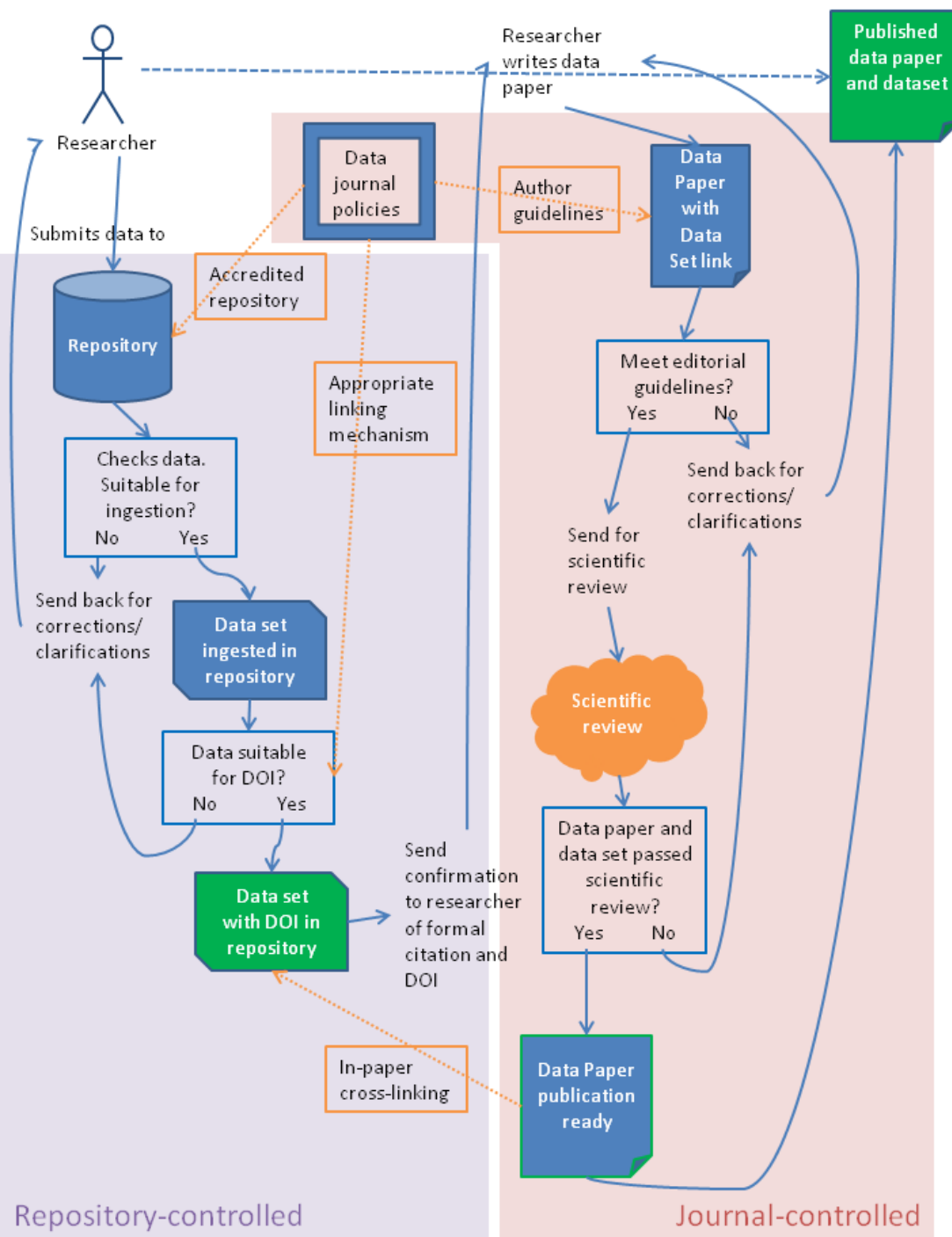
- *Developing Data Attribution and Citation Practices and Standards* workshop, Berkeley, California August 2011
http://www.nap.edu/catalog.php?record_id=13564
- *Data Citation Principles* workshop, IQSS, Harvard University, Massachusetts, May 16-17, 2011
- literature survey and interviews with members of stakeholder communities (researchers, funders, data repository managers).



Emerging Principles for Data Citation (2)

1. First Class Status
2. Persistence
3. Granularity
4. Resolvability
5. Attribution
6. Metadata Standards

Principles taken from:
TG on Data Citation Report Outline
Presentation to Roundtable Discussion
Taipei, Taiwan October 2012

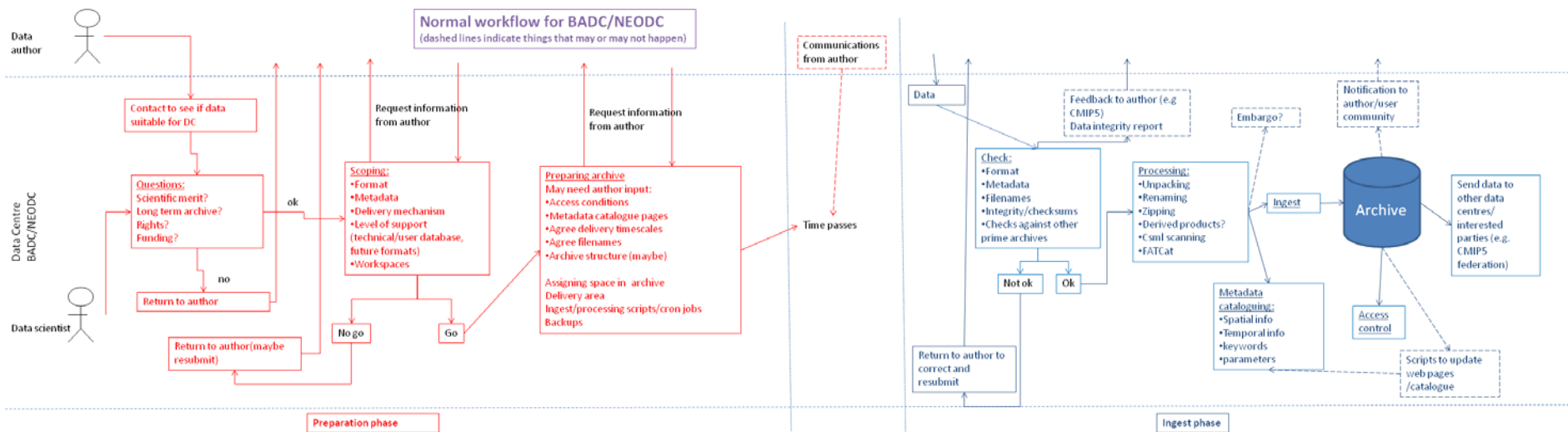


Example of (potential) steps/workflow required for a researcher to publish a data paper

- Items in **orange** refer to the areas of interest which will be investigated in PREPARDE
- Division of area of responsibilities between
 - *repository controlled* processes
 - *journal controlled* processes

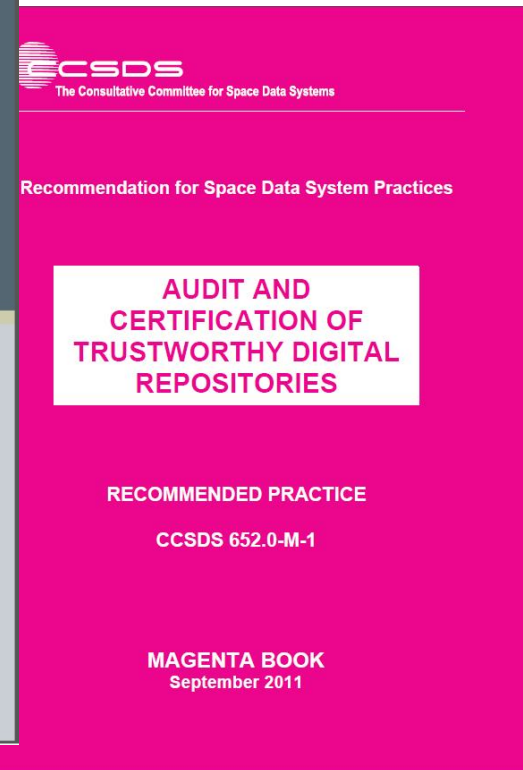
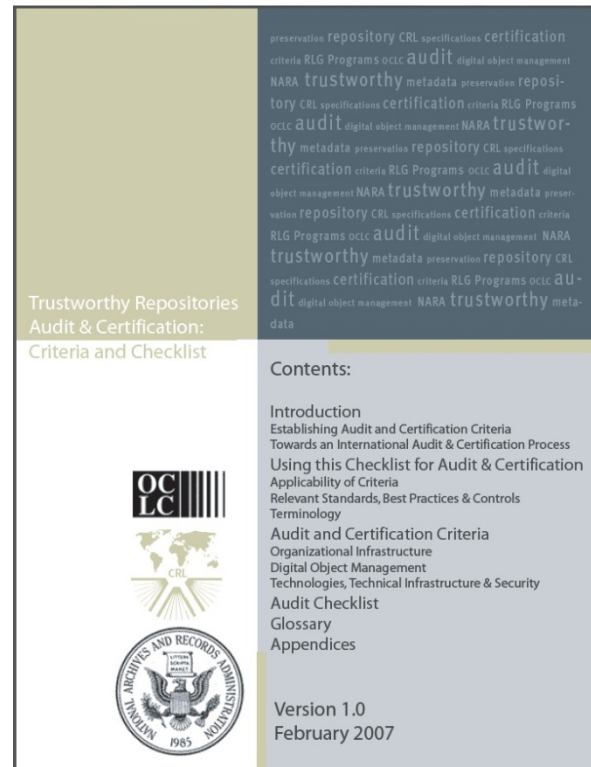
Data repository workflows

- Data centre and journal workflows captured
- Work on comparisons and identification of cross-linking points is continuing.
- Workshop planned for 2013 – if interested please email sarah.callaghan@stfc.ac.uk

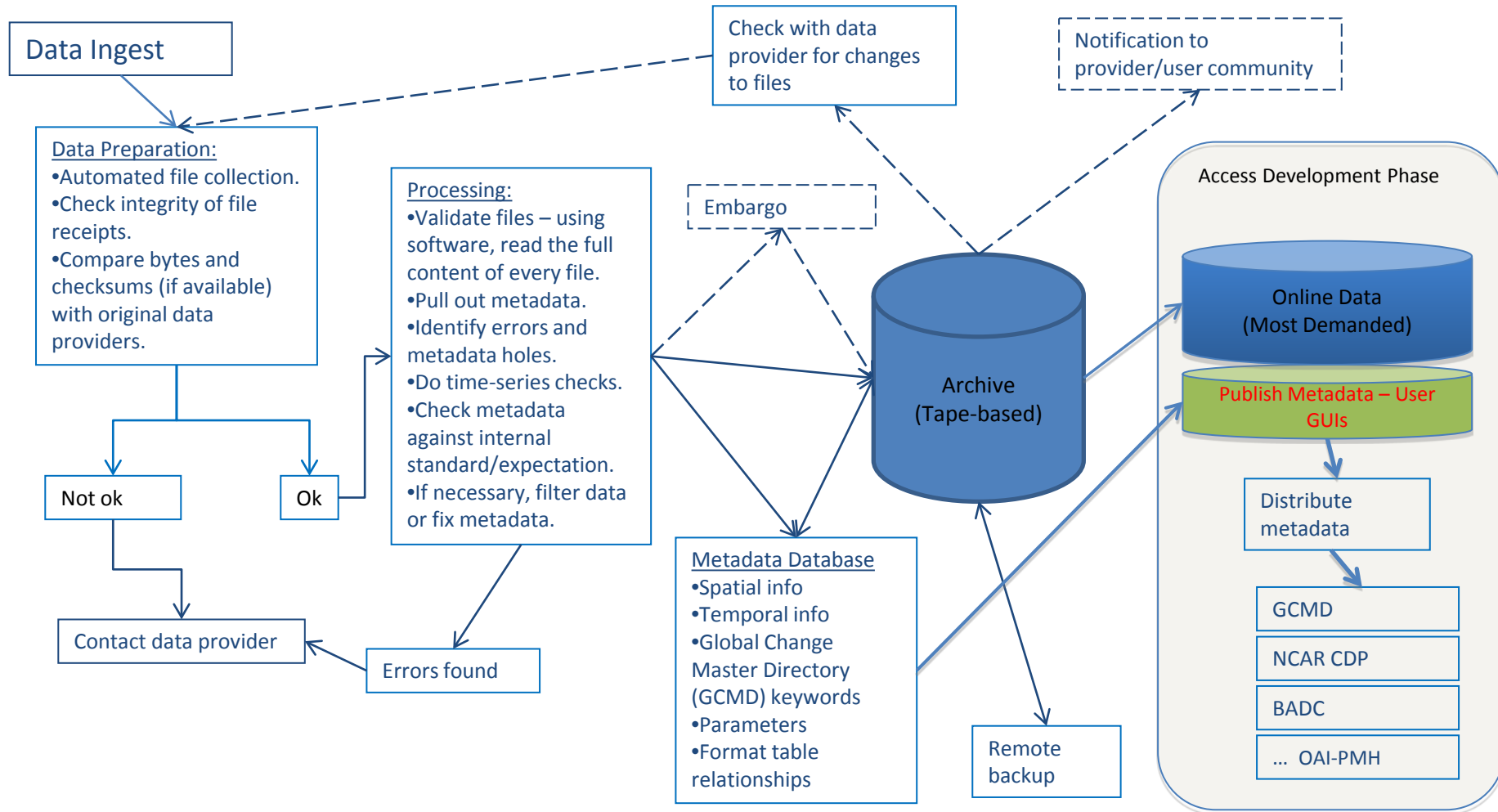


Repository Accreditation & Workflows

- Repository certification procedures focus on repository “trustworthiness”
- Assessing trustworthiness requires assessing the entire repository workflow



Repository Workflow – NCAR Comp. & Info. Systems Lab Research Data Archive (RDA)



Repository Workflow – NCAR Earth Observing Lab Data Management Group

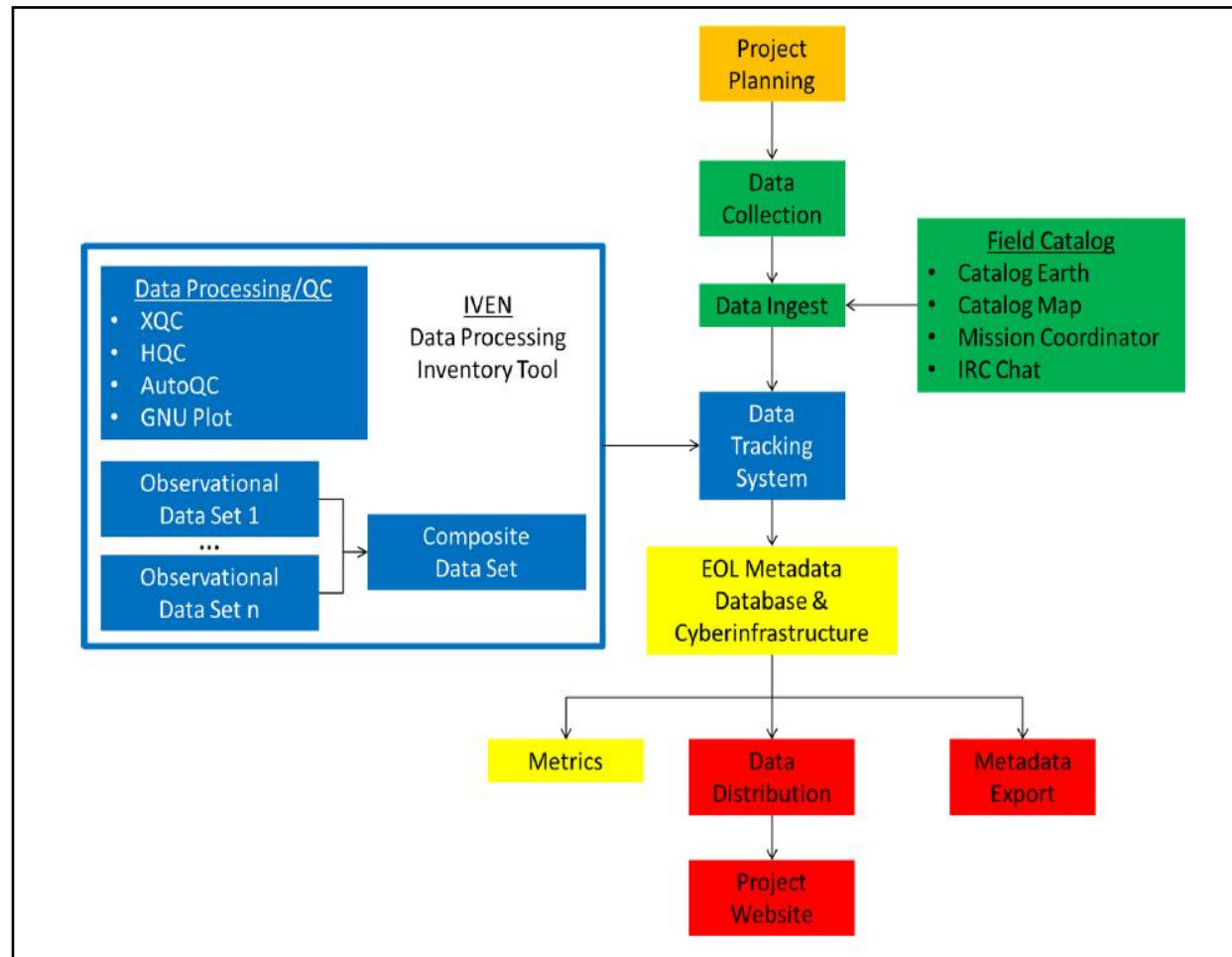


Diagram by Christopher Eaker

Workflows

- Data quality assessment processes take place throughout the workflow
- Data publication is often not a one-time process
 - Data might be used in research publications before being made publicly available.
 - Data might be updated or fixed due to user feedback after being posted.

Repository Accreditation

- What makes a repository trustworthy?
 - Many things: mission, processes, expertise, workflows, history, systems, documentation, ...
 - Assessing trustworthiness is a complex and holistic task
- Peer review of data is implicitly peer review of repository
- IDCC workshop #IDCC13 - 17 Jan 2013, Amsterdam
 - Report to follow

Discussion!

