



Certification as a means of providing trust: the European framework

Ingrid Dillo

Data Archiving and Networked Services

Data Archiving and Networked Services

DANS

Trust and Digital Preservation, Dublin June 4, 2013

Certification as a means of providing trust: the European Framework

- Definitions research data
- Data sharing and trust
- European Framework
- The devil's advocate



What is DANS?

Institute of Dutch
Academy and
Research Funding
Organisation
(KNAW & NWO)
since 2005

**Mission: promote
permanent access
to digital research
information**

First predecessor
dates back to
1964 (Steinmetz
Foundation),
Historical Data
Archive 1989

DANS

main activities and services



- Encourage researchers to self-archive and reuse data by means of our Electronic Archiving SYstem **EASY**
- Our largest digital collections are in archaeology, social sciences and history (moving into other domains)
- Provide access, through narcis.nl, to thousands of scientific datasets, e-publications and other research information in the Netherlands
- Data projects in collaboration with research communities and partner organisations
- Advice, training and support (Data Seal of Approval, Persistent Identifier Infrastructure)

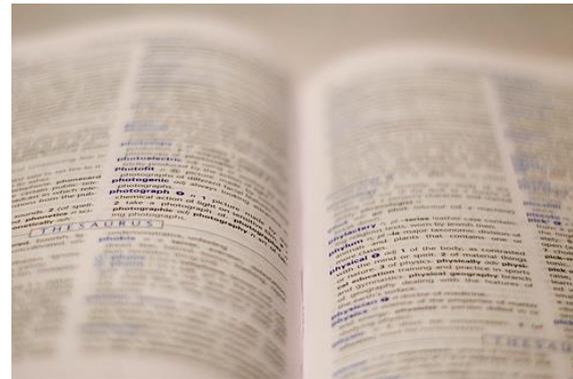
Definition research data



Factual records, which may take the form of numbers, symbols, text, images or sounds, used as primary sources for research, commonly accepted in the research community as necessary to validate research findings.

Definition research data

Data in any format or medium that relate to or support research, scholarship, or artistic activity.



Classified as:

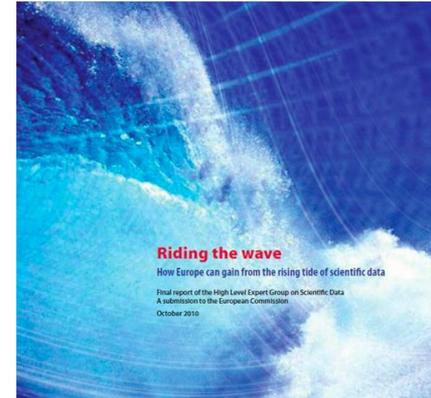
- **Raw or primary data:** Information recorded as notes, images, video footage, paper surveys, computer files, etc.
- **Processed data:** Analyses, descriptions, and conclusions prepared as reports or papers
- **Published data:** Information distributed to people beyond those involved in data acquisition and administration





Data is hot!

- Riding the wave: *critical importance of sharing and preserving reliable data produced during the scientific process*
- Commissioner Kroes: “*Data is the new gold*”
- EC Recommendation on access to and preservation of scientific information:
European open access policy



Peter Middendorp: 'Ik zag liefdeswalging in haar ogen, maar ze hield zich groot' sp12

de Volkskrant

MAANDAG 29 APRIL 2013
VOLKSKRANT.NL

DE VOLKSKRANT BIJ JACOB ROEMERPLAATS 5, POSTBUS 1000, 1000 BA, AMSTERDAM | REDACTIEVOLKSKRANT.NL | REDACTIE 020-632 9322 | KLANTENSERVICE 088-06 1161 | BEZORGING 020-634 1510



Hoe heeft Amsterdam zich voorbereid?

22 SCENARIO'S EN HET ORANJEVOEL IN DE ETALAGES PA-7



Gouden afscheid

SAAI IS EDITH BOSCH HOOFD GEWEEST SP10-11

illegaali
Extra ledenraad PvdA

Data fraud: front page news on 29/4/2013

De waarheid volgens Diederik Stapel

Voor het eerst sinds zijn fraude aan het licht kwam, spreekt de psycholoog. 'Ik had geen stopknop.'

PAGINA V2-4



Foto: Kees Broekman

Wetenschap

INTERVIEW DIEDERIK STAPEL

Voor het eerst sinds zijn publieke val liet Diederik Stapel zich interviewen. The New York Times sprak niet alleen de fraudente ex-hoogleraar, maar ook diens vrouw, ouders en ex-bondgenoten.

Door Yudhijit Bhattacharjee

Het brein van een bedrieger

Op een onwaarschijnlijk moment in de jaren vijftig van de vorige eeuw werd een groep van psychologen gevraagd om te onderzoeken hoe de hersenen van bedriegers werken. De onderzoekers ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken. Ze ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken. Ze ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken.

De onderzoekers ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken. Ze ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken. Ze ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken.

De onderzoekers ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken. Ze ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken. Ze ontdekten dat bedriegers een ander patroon vertoonden in de manier waarop ze hun hersenen inzetten om te denken.



De fraude van Diederik Stapel leidde tot een enorme hausse aan media aandacht

'Ik merk dat hij steeds als een kammerlid uitlijt'
Stapel, zegt dat ik de deur moet dichtdoen. Hij wekt dat hij zijn leven lang een symbool is geweest van de wetenschap.

Stapel, zegt dat ik de deur moet dichtdoen. Hij wekt dat hij zijn leven lang een symbool is geweest van de wetenschap.

Stapel, zegt dat ik de deur moet dichtdoen. Hij wekt dat hij zijn leven lang een symbool is geweest van de wetenschap.

DE VOLKSKRANT
MAANDAG 29 APRIL 2013

V3

CV Diederik Stapel

1964	1965	1966	1967	1968	1969
1970	1971	1972	1973	1974	1975
1976	1977	1978	1979	1980	1981
1982	1983	1984	1985	1986	1987
1988	1989	1990	1991	1992	1993
1994	1995	1996	1997	1998	1999
2000	2001	2002	2003	2004	2005
2006	2007	2008	2009	2010	2011
2012	2013	2014	2015	2016	2017



Why is that?

- **Transparency and replication of research**
- Hot item in the Netherlands



News

Report finds massive fraud at Dutch universities

Investigation claims dozens of social-psychology papers contain faked data.



Science Insider

Breaking news and analysis from the world of science policy

Report: Dutch 'Lord of the Data' Forged Dozens of Studies (UPDATE)

by Gretchen Vogel on 31 October 2011, 7:05 PM | [34 Comments](#)

theguardian

Public sector can combat fraud with data sharing

Outsourcing is not the only thing to blame for procurement fraud, says **Graham Kemp**, and the public sector needs to view data less as a security risk but knowledge to be shared

Fraud Case Seen as a Red Flag for Psychology Research

By BENEDICT CAREY

Published: November 2, 2011

A well-known psychologist in the Netherlands whose work has been published widely in professional journals falsified data and made up entire experiments, an investigating committee has found. Experts say the case exposes deep flaws in the way science is done in a field, [psychology](#), that has only recently earned a fragile respectability.



SPIEGEL ONLINE

Niederlande

Renommierter Psychologe gesteht Fälschungen



PLOS one

accelerating the publication of peer-reviewed science

Search an

[Home](#) [Browse Articles](#) [About](#) [For Readers](#) [For Authors and Reviewers](#)

RESEARCH ARTICLE

OPEN ACCESS

How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data

Why is that?

- **Re-use of data**

- Efficiency in research
- Combining datasets
- Usage of datasets in other disciplines
- Usage of data outside academia
- Higher return on investment



Importance of research data

Data



Information



Presentation



Knowledge



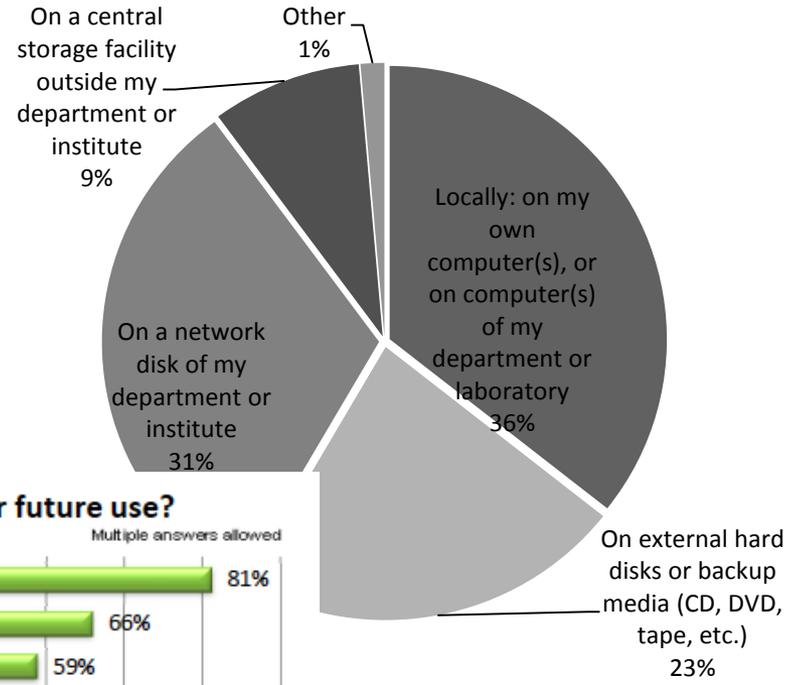
Reality of data sharing



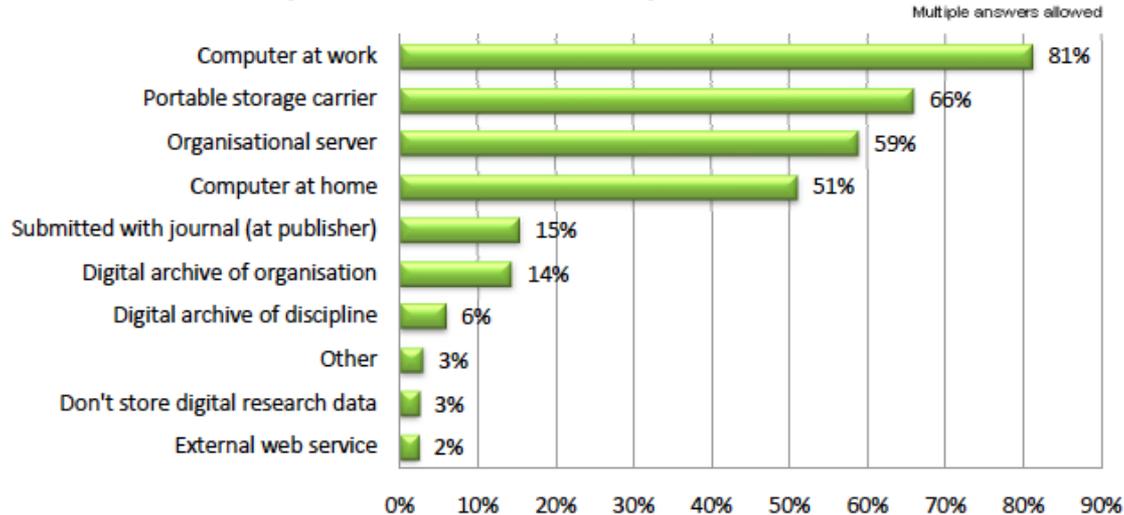
Insight into digital preservation of research output in Europe



survey report

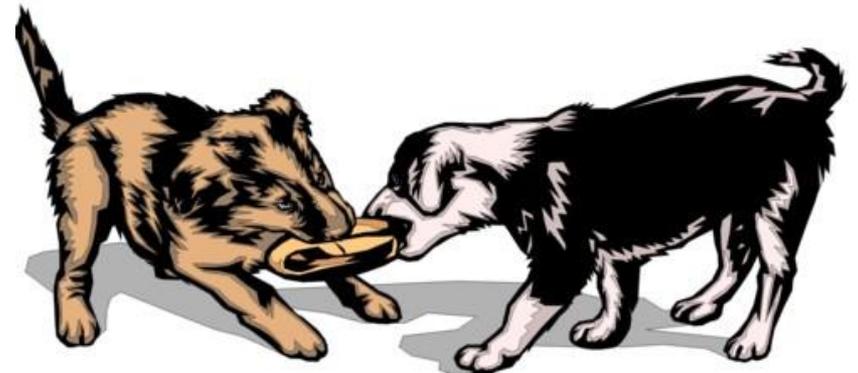


Where do you as a researcher store your data for future use?



Why not share?

- Those data are mine!
- Discredit my findings
- Still analyzing the data
- I cannot **trust** the data produced somewhere else



Trust

Trust is at the very heart of storing and sharing data

- Users
- Depositors
- Funders



Trust: What do we mean? What do we rely on?



RANG IS ALLEEN RANG ALS ER RANG OP STAAT

RANG
TONNEMA N.V. SNEEK

CRIS
FRU

'n produkt van de KINGFABRIEKEN TONNEMA N.V. SNEEK

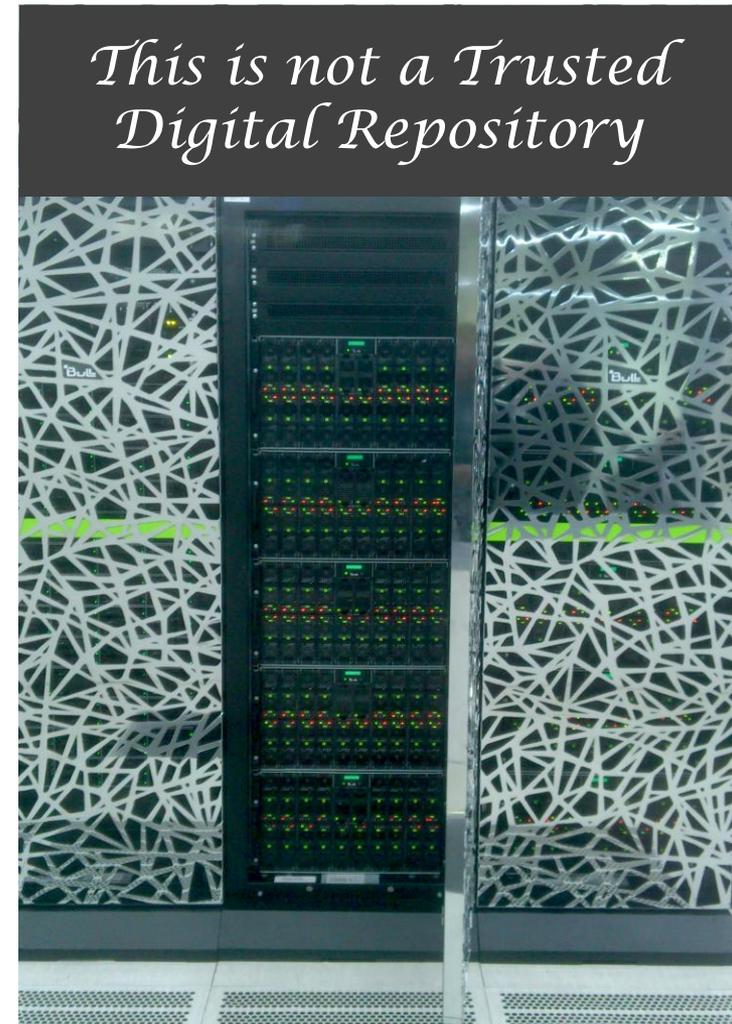
An advertisement for RANG candy. It features a blue tube of candy with the word 'RANG' in large white letters and 'TONNEMA N.V. SNEEK' in smaller white letters. The tube is tilted and shows several pieces of candy inside. Below the tube are several white flowers with red and yellow centers. The background is a light blue sky. The text 'RANG IS ALLEEN RANG ALS ER RANG OP STAAT' is written in white on a blue background at the top left. At the bottom, it says ''n produkt van de KINGFABRIEKEN TONNEMA N.V. SNEEK'.

What is a trustworthy digital repository?



Things are not always what they say they are.

Things do not always state what they are.



What is trust built on (in trusted archives)?

- Dedicate yourself (mission statement)
- Do what you promise (stable, sincere and competent reputation)
- Be transparent (peer review, get certified)



"They don't trust each other to share research."



What is a trustworthy digital repository?

- mission to provide reliable, long-term access to managed digital resources to its designated community, now and into the future
- constant monitoring, planning, and maintenance
- understand threats to and risks within its systems
- regular cycle of audit and/or certification



Standards of trust

THE UK'S LARGEST COLLECTION OF DIGITAL RESEARCH DATA IN THE SOCIAL SCIENCES AND HUMANITIES

UK • DATA
ARCHIVE

HOME

ABOUT US

CREATE &
MANAGE DATA

DEPOSIT
DATA

HOW WE
CURATE DATA

HOW TO CURATE DATA

STANDARDS OF TRUST

HOW WE CURATE DATA

THE PROCESS

OUR QUALITY CONTROL

OUR PRESERVATION
POLICY

TRUSTED DIGITAL
REPOSITORIES

STANDARDS OF TRUST / OVERVIEW

OVERVIEW

DATA SEAL OF APPROVAL

ISO16363

DIN 31644

Any organisation which provides access to data over a long period of time should be fully trusted only with a public statement describing the practices they follow and the provenance of data they provide. Standards of trust are critical.

Assessing trustworthiness



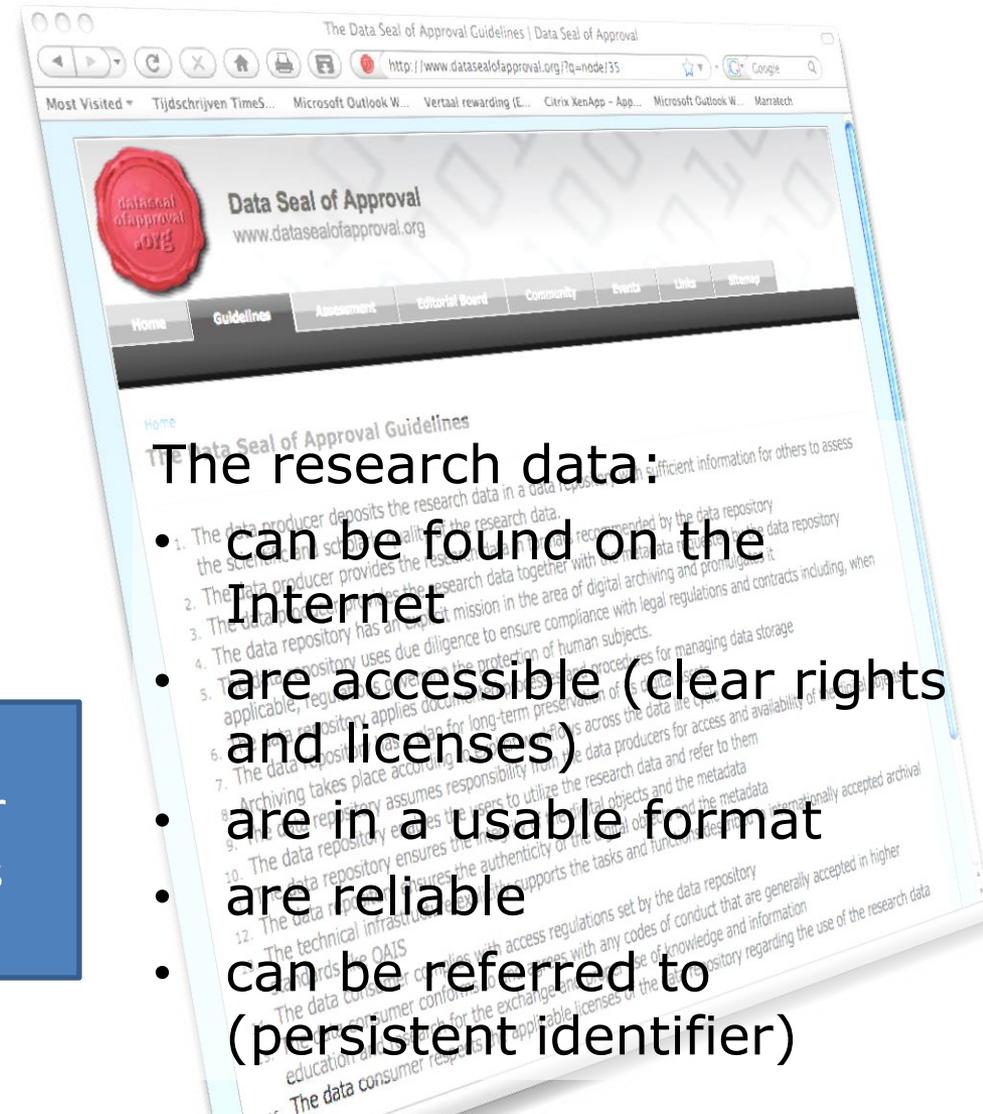
DIN



Certification Standards: Data Seal of Approval (DSA)

- DANS initiative
- International Board
- 16 guidelines
- Self assessment
- Transparency

Data producers are responsible for the quality of research data, repositories for storage and long-term access, and users for correct use of data



Certification Standards: NESTOR - DIN

- Germany
- Competence network for digital preservation
- WG Certification



Certification Standards: DIN 31644

- 34 criteria
- Test audits 2013



nestor-Kriterien
Kriterienkatalog
vertrauenswürdige
digitale Langzeitarchive

The logo for DIN (Deutsches Institut für Normung) is displayed in a bold, black, sans-serif font. The letters "DIN" are centered between two thick, horizontal black bars.

Certification Standards: TRAC -> ISO

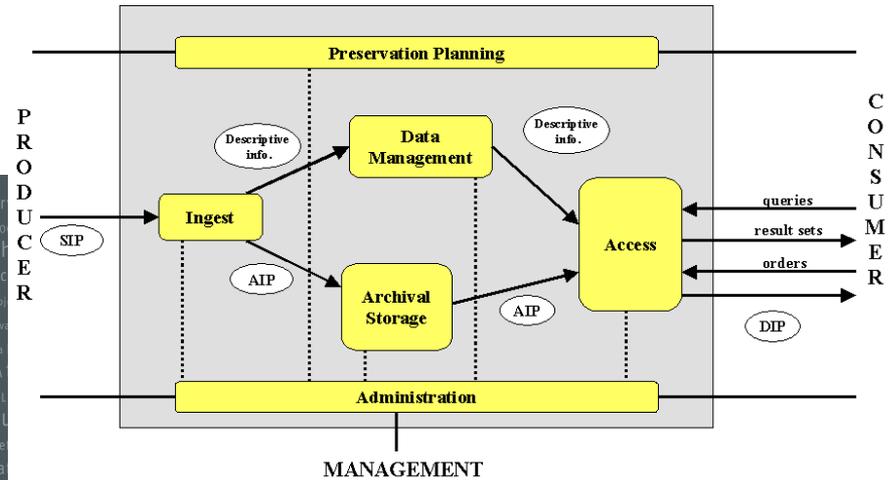
- OAIS-model
- TRAC
- ISO

Trustworthy Repositories
Audit & Certification:
Criteria and Checklist

Contents:

- Introduction
 - Establishing Audit and Certification Criteria
 - Towards an International Audit & Certification Process
- Using this Checklist for Audit & Certification
 - Applicability of Criteria
 - Relevant Standards, Best Practices & Controls
 - Terminology
- Audit and Certification Criteria
 - Organizational Infrastructure
 - Digital Object Management
 - Technologies, Technical Infrastructure & Security
- Audit Checklist
- Glossary
- Appendices

Version 1.0
February 2007

Certification Standards: ISO 16363

- Over 100 metrics
- Test audits 2011
- Full external auditing process



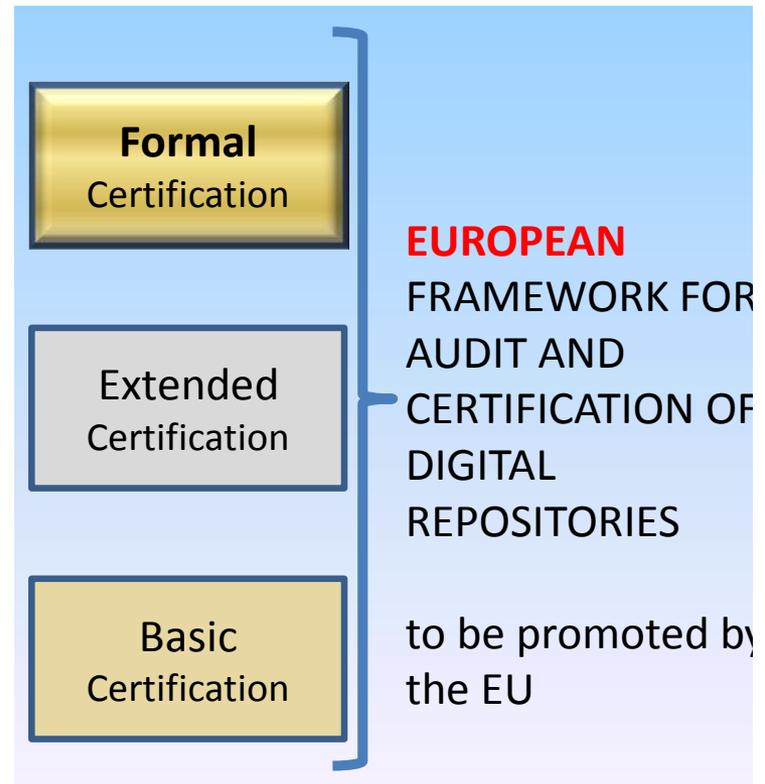
Certification Standards: ISO metrics

- Organisational infrastructure
- Digital object management
- Infrastructure and security risk management
- Metric
 - Supporting text
 - Examples of evidence
 - discussion



European Certification Framework

- MoU
- Three-tiered framework



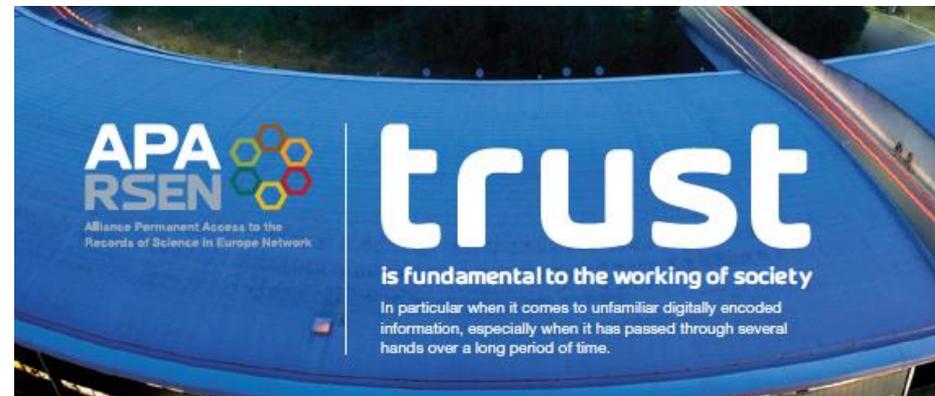
Framework levels

- **Basic Certification** is granted to repositories which obtain DSA certification
- **Extended Certification** is granted to Basic Certification repositories which *in addition* perform a structured, externally reviewed and publicly available self-audit based on ISO 16363 or DIN 31644
- **Formal Certification** is granted to repositories which *in addition to* Basic Certification obtain full external audit and certification based on ISO 16363 or DIN



On-going work on Trust

- Work Package on “Trust” within APARSEN project
- European Framework for Audit and Certification of Digital Repositories – seek collaboration with ICSU World Data System (accreditation)
- RDA Group on Certification



The devil's advocate



- Trustworthiness of digital repositories is an illusion
 - Too complicated to measure
 - Impossible to maintain over time
 - Too informal – too expensive and time consuming



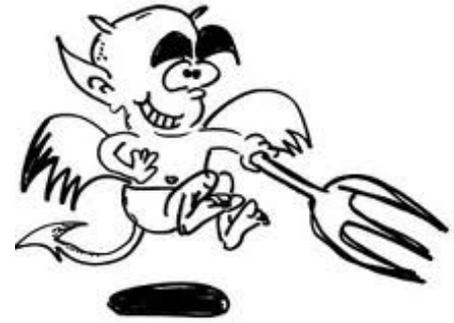
My message



- Trustworthiness of digital repositories is not an illusion
 - Three levels are clear and balanced
 - Different levels for different needs



The devil's advocate



- Objective and consistent auditing is an illusion
 - “If auditing becomes a career, what will happen to objectivity?” (H. Tibbo)
 - Impossible to guarantee consistency around the globe



My message



- Objective and consistent auditing can be done
 - Auditing already is a career
 - Do not stress consistency around the globe too much



What does *trust* mean?

- Trust by definition implies uncertainty
- “You don't have to see the whole staircase, just take the first step.”
(Martin Luther King)





Thank you for your attention!

ingrid.dillo@dans.knaw.nl

<http://datasealofapproval.org/>

www.trusteddigitalrepository.eu

make
the digital world
more
reliable