Stefanie Mühlhausen, Kevin Lüdemann, Lars Quentin

Best Practices for Administrators

How to make RTFM and RTFD your standard responses

Table of contents

1 Introduction

2 Different Frameworks

3 Documentation

4 Data protection

What is a best practice?

“A best practice is a method or technique that has been generally accepted as superior to other known alternatives because it often produces results that are superior to those achieved by other means or because it has become a standard way of doing things, e.g., a standard way of complying with legal or ethical requirements.”

— https://en.wikipedia.org/wiki/Best_practice

Why do we need best practices?

- Why re-invent the wheel, if there is a *working* solution at hand?
- Ease transitions and communication in teams
- Save you a lot of time in the long run

Where do best practices come from?

Depending on the field, you will find different requirements for best practices:

- Project Management
  - ISO 21500, ISO 9000, PRINCE2, HERMES (Switzerland), PM^2 (EU)
- Software Development / Software Lifecycle Management
  - Waterfall or Spiral Model, Scrum, Kanban
- Quality Management
  - ISO 9000, ISO 9001, (ITIL)
- Information Security Management (Systems)
  - ISO 27001, ITIL
- IT-Service Management
  - ITIL, COBIT, eTOM, FitSM, ISO 20000

Information Technology Infrastructure Library (ITIL)

- Set of detailed practices for IT activities (e.g., ITSM and ITAM)
  - incident management
  - asset management
  - event management
  - IT security management
  - change management
  - service management
- De-facto-standard $\rightarrow$ certification for individuals
- Describes processes, procedures, tasks, checklists

It’s all about the documentation

Different types of documentation necessary

- User documentation
- Network documentation
- Software documentation
- Hardware documentation
- Process documentation
- Change documentation (logging)

What and how you will need to document, depends on your systems, your team, your employer and your responsibilities

→ **Documentation is extremely individual.**

Attention to the audience!

User-faced

- Different levels of expertise
- Mainly focussed on how to *use* the system
  - Tutorials and how-tos
  - FAQ
  - Code examples

Attention to the audience!

User-faced

- Different levels of expertise
- Mainly focussed on how to *use* the system
  - Tutorials and how-tos
  - FAQ
  - Code examples

Admin-faced

- Documents workflows and responsibilities
- Explains how to *run* the system
  - Where to find what
  - How to do what
  - What the admin did and **why**

Attention to the audience!

User-faced

- Different levels of expertise
- Mainly focussed on how to *use* the system
  - Tutorials and how-tos
  - FAQ
  - Code examples

This saves you from spending too much time on easy user support questions and makes the customer happy. (=RTFM)

Admin-faced

- Documents workflows and responsibilities
- Explains how to *run* the system
  - Where to find what
  - How to do what
  - What the admin did and **why**

This saves you from lengthy training of new team members and from calls from your team mates during your vacation. (=RTFD)

Exercise 1 Recall and document - 10 minutes

Recall what you have set up and installed yesterday. Document your steps from your memory. Create admin-faced documentation on your Warewulf setup.

- Do you remember every step?
- Are you having difficulties remembering everything?
- Could you explain to a colleague what exactly he*she has to do?

For the rest of the course, write documentation as you go!

Requirements for good documentation (1/2)

- Write documentation that is **inviting** and clear

  Avoid assumptions about what others may know
- Write documentation that is **comprehensive**

  Fully document all features, exceptions etc.
- Write documentation that is **skimmable**

  Use headlines, bulleted lists, table of content to ease navigation
- Write documentation that offers **examples** of how to use the software

  Illustrate the most common use case

https://www.oreilly.com/content/the-eight-rules-of-good-documentation/

Requirements for good documentation (2/2)

- Write documentation that has **repetition**, when useful

  *Code may be DRY, documentation ARID (**a**ccept (some) **r**epetition **i**n **d**ocumentation)*
- Write documentation that is **up-to-date**

  Incorrect documentation is worse than missing documentation
- Write documentation that is **easy to contribute** to

  Use software development tools and techniques for documentation
- Write documentation that is **easy to find**

  Keep an updated README, link to the more extensive documentation

https://www.oreilly.com/content/the-eight-rules-of-good-documentation/

What is good / bad documentation and why?

Good Documentation

Example: A job submission guide with a working script, error troubleshooting, and links to resources.

Bad Documentation

Example: “Run the script” — no path, no arguments, no error handling.

Why it matters

- **Good** → Saves time, reduces errors, enables onboarding
- **Bad** → Causes frustration, leads to repeated questions, increases risk

Exercise 2 Peer Review of Documentation - 10 Minutes

In your group, compare the documentation you created during the previous exercise. Discuss and agree on:

- One strong aspect (e.g., clear structure, helpful example, good audience fit)
- One area for improvement (e.g., missing context, unclear steps, hard to navigate)

Which tool to use is extremely individual

Questions when choosing a documentation tool include:

- What is your budget, in terms of financial cost and time (development, maintenance, training)
- What kind of output do you need? Web, PDFs, support tools, code base
- Skills/ tech comfort (WYSIWYG, Markdown, ...)
- Collaboration workflow & Integration with Existing Tools (Confluence, Git, ...)
- Ease of use & Long-term maintainability
- Do you need to support multiple languages?

Optional exercise 3 Deep Dive into ITIL

Read into the ITIL 4 Foundation. Focus on:

- Service Value System
- The Four Dimensions of Service Management

Next, research the relevance of two of the four dimensions for HPC system administration.

Data Protection

- In the EU, it is defined by the *General Data Protection Regulation* (GDPR)
  - Datenschutzgrundverordnung
- It defines the regulatory policies and requirements on data protection
- GDPR contains over 80 opening clauses, allowing states to decide further regulation.
  - Bundesdatenschutzgesetz (BDSG)
  - Niedersächsisches Datenschutzgesetz (NDSG)

Data Protection

- In the EU, it is defined by the *General Data Protection Regulation* (GDPR)
  - Datenschutzgrundverordnung
- It defines the regulatory policies and requirements on data protection
- GDPR contains over 80 opening clauses, allowing states to decide further regulation.
  - Bundesdatenschutzgesetz (BDSG)
  - Niedersächsisches Datenschutzgesetz (NDSG)

Definitions (based on Art 4 GDPR)

- **Personal Data:** Any information relating to an identified *or identifiable* natural person
  - So if its POSSIBLE to identify!
- **Processing:** Any operation [...] which is performed on personal data [...], whether or not by automated means
- **Personal Data Breach:** Breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data [...]

Overview GDPR: The relevant parts

- Principles for Processing (Art 5)
- What the Controller has to do (Art 24)
  - Data protection by design and by default (Art 25)
  - Processor / Auftragsverarbeiter (Art 28)
  - Technical and Organizational Measures (Art 32)
  - Records of processing activities / VVT (Art 30)
  - Data Protection Impact Assessment / DSFA (Art 35+36)
- Rights of the data subject (Art 12-23, Art 33+34)
  - Transparent Information
  - Right to Rectification, Erasure (’right to be forgotten’), Restriction of Processing, Data Portability, Objection
  - Notification obligation towards data subject and authority
- Supported by your Data Protection Officer (Art 37-39)

Principles of Processing (Art 5 GDPR)

- **Lawfulness, Fairness and Transparency**
  - Transparently document, not misleading
- **Purpose Limitation**
  - Data only for agreed-upon use
- **Data Minimisation**
  - Only collect as much as you need
- **Accuracy**
  - accurate and, where necessary, kept up to date (as reasonably possible)
- **Storage Limitation**
  - Only store as long as needed
- **Integrity and Confidentiality**
  - “processed in a manner that ensures appropriate security of the personal data”
- **Accountability**
  - Document what you did for the other 5

Technical and Organisational Measures (TOMs)

- Data has to be secured using both technical and organisational measures
- GDPR follows a **risk based** approach:
  The higher the risk, the more secure it needs to be
  - Risk $=$ Likelyhood $\cdot$ Severity
- Taking into account the **“State of The Art”** / “Stand der Technik”
- Question: What is the State of the Art?

Technical and Organisational Measures (TOMs)

- Data has to be secured using both technical and organisational measures
- GDPR follows a **risk based** approach:
  The higher the risk, the more secure it needs to be
  - Risk = Likelyhood · Severity
- Taking into account the **“State of The Art”** / “Stand der Technik”
- Question: What is the State of the Art?
- Answer: Whatever the State of the Art Document says (no joke)
  - Teletrust: Handreichung Stand der Technik in der IT-Sicherheit
  - BSI IT-Grundschutz 200-2
    - Explainations of it in the IT-Grundschutz- Kompendium

Applied Data Protection: What can you do

- The Standardatenschutzmodell translates the GDPR into 7 actionable goals

1. **Data Minimization:** Collecting less data, pseudonymisation
2. **Availability:** RAID, georedundancy, monitoring
3. **Integrity:** Backups, signing, RAID
4. **Confidentiality:** Encryption, RBAC, segmentation, multitenancy
5. **Unlinkability:** No shared databases, Principle of least privilege
6. **Transparency:** User documentation, privacy policy
7. **Intervenability:** Standardized Workflows, RAG instead of fine tuning

Exercise 4 Data Protection in HPC - 10 Minutes

Apply the 7 Goals of the Standard Data Protection Model (SDM) to the operation of an HPC cluster. Identify relevant TOMs (Technical, Organizational, Management measures) and assess potential risks.

Discuss your findings in your group.

Plenary discussion

- Open questions?
- What were your findings?
- What would you do differently?

Reminder: **For the rest of the course, write documentation as you go!**

https://chat-ai.academiccloud.de/chat?settings=
eyJzeXN0ZW1fcHJvbXB0IjoiWW91IGFyZSBhbiBleHBlcnQgdGVjaG5pY2FsIHdyaXRlciBz