FAIR and ALCOAplus
data management principals

A comparison of two data documentation principles

PAASP

FAIR

Findable
The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the **FAIRification process**.
>   F1. (meta)data are assigned a globally unique and eternally persistent identifier.
>   F2. data are described with rich metadata.
>   F3. (meta)data are registered or indexed in a searchable resource.
>   F4. metadata specify the data identifier.

Accessible
Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.
>   A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
>   A1.1 the protocol is open, free, and universally implementable.
>   A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
>   A2 metadata are accessible, even when the data are no longer available.

Interoperable
The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.
>   I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
>   I2. (meta)data use vocabularies that follow FAIR principles.
>   I3. (meta)data include qualified references to other (meta)data.

Reusable
The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.
>   R1. meta(data) have a plurality of accurate and relevant attributes.
>   R1.1. (meta)data are released with a clear and accessible data usage license.
>   R1.2. (meta)data are associated with their provenance.
>   R1.3. (meta)data meet domain-relevant community standards.

The principles refer to three types of entities: data (or any digital object), metadata (information about that digital object), and infrastructure. For instance, principle F4 defines that both metadata and data are registered or indexed in a searchable resource (the infrastructure component).

source

PAASP

ALCOA(+)

The industry standard established by the WHO with the initialism known as ALCOA, which has been expanded on to ALCOA+ (currently used by the FDA, WHO, PIC/S and GAMP – Data Integrity), neatly requires **all data** to have the following qualities:

- **Attributable** — Who acquired the data or performed an action and when?
- **Legible** — Can you read the data and any entries?
- **Contemporaneous** — Was it recorded as it happened?
- **Original** — Is it the first place data is recorded?
- **Accurate** — Are all the details correct?

- **Complete** — Are all data included (any repeat or reanalysis performed on the sample)?
- **Consistent** — Are all elements in chronological order?
- **Enduring** — Are all recordings and notes accessible over extended period?
- **Available** — Can the data be accessed for review over the lifetime of the record?

Based on these attributes being adhered to the data can be trusted, this becomes both simpler and more complicated when we introduce electronic systems capable of managing all these attributes as part of the 'meta-data' and have to consider how we can retrieve, store and archive data across it's entire life cycle.

source

PAASP

FAIR and ALCOA

FAIR focuses on

- Dealing with data
- Meta-data
- Data infrastructure

Origin

The FAIR Data Principles (Findable, Accessible, Interoperable, Reusable) were drafted at a Lorentz Center workshop in Leiden in the Netherlands in 2015.

Nature article for launch

ALCOA focuses on

- DATA INTEGRITY
- Context of data
- Content of data

Origin

WHO guidance, Section 9 Good Documentation Practice (LINK)

PAASP

WHO guidance – original text

9. Good documentation practices

9.1 The basic building blocks of good GXP data are to follow GDocP and then to manage risks to the accuracy, completeness, consistency and reliability of the data throughout their entire period of usefulness – that is, throughout the data life cycle.

Personnel should follow GDocP for both paper records and electronic records in order to assure data integrity. These principles require that documentation has the characteristics of being attributable, legible, contemporaneously recorded, original and accurate (sometimes referred to as ALCOA). These essential characteristics apply equally for both paper and electronic records.

9.2 **Attributable.** Attributable means information is captured in the record so that it is uniquely identified as executed by the originator of the data (e.g. a person or a computer system).

9.3 **Legible, traceable and permanent.** The terms legible and traceable and permanent refer to the requirements that data are readable, understandable, and allow a clear picture of the sequencing of steps or events in the record so that all GXP activities conducted can be fully reconstructed by the people reviewing these records at any point during the records retention period set by the applicable GXP.

9.4 **Contemporaneous.** Contemporaneous data are data recorded at the time they are generated or observed.

9.5 **Original.** Original data include the first or source capture of data or information and all subsequent data required to fully reconstruct the conduct of the GXP activity. The GXP requirements for original data include the following:

   •original data should be reviewed;
   •original data and/or true and verified copies that preserve the
content and meaning of the original data should be retained;
   as such, original records should be complete, enduring and readily retrievable and readable throughout the records retention period.

9.6 **Accurate.** The term "accurate" means data are correct, truthful, complete, valid and reliable.

9.7 Implicit in the above-listed requirements for ALCOA are that the records should be **complete, consistent, enduring and available** (to emphasize these requirements, this is sometimes referred to as ALCOA-plus).

9.8 Further guidance to aid understanding as to how these requirements apply in each case and the special risk considerations that may need to be taken into account during implementation are provided in Appendix 1.

PAASP