OPEN DATA & DATA CLEANING

Anton Van de Putte

FAIR

a set of 15 guiding principles

encourages comprehensive & sustainable data management

involves “everyone”: producers, stewards and users

FAIR

FAIR IS NOT

not a standard

not equal to ‘Open’ or ‘Free’

no guarantee for successful (open) science

Make a plan for the data before you start a project!

Collect detailed descriptive information (= metadata) throughout

Use standards and formats common to your discipline

Store the data in a trusted & sustainable repository or data center

See to that the data gets persistent identifiers (DOIs)

Apply a suitable usage license

Provide end users with information on “intended use”

Make the data “as open as possible, as closed as necessary”

Ensure that metadata remain available even if the data cannot be accessed any more

FAIRness needs to be applied where it makes sense

OUR
PRECIOUS
DATA

THE FIRST USER OF THIS DATA IS FUTURE **YOURSELF**

BE LAZY, DO IT RIGHT THE FIRST TIME

WHAT

WHERE

WHEN

WHO

HOW

DARWIN CORE

terms in the Darwin Core quick reference guide.

REQUIRED TERMS

basisOfRecord	occurrenceID	scientificName	scientificNameID
occurrenceStatus	decimalLongitude	decimalLatitude	eventDate