# How to organize your harddisk?

Why is it easier to find something on the Internet than on your hard disk?

#### On folders and files, not data sets

- General talk on folder and file naming and organization on harddisk
- Formal Data set organization and naming covered by Reinhard in another talk

## Common problems with data on hard disk

- User cannot easily find information again, and spend time looking for it. 80% of users fail the "show me" test
- File and folder names inconsistent and lack clear structure
- Different types of data (private, public, shared, nonshared) is mixed, and difficult to share data
- Different versions is mixed up, and difficult to find final/last version
- Data is insufficiently documented, and other people cannot understand it (and after a while neither can the original researcher)

## What is the problem of the hard disk?

- We are forced to use simple tree-structure, with folders inside folders, and the search facilities are simple
- But real world logic multi-dimensional for example, topic, organizational context, time, relation to other steps in process
- Internet pages are tagged with meta data, linked, and indexed in very sophisticated ways

## Organizing principles for information

| Organizing principle   | Pros-cons   |
|--|---|
| Organization, people, projects                                       | Subject to change, but legal necessity (owner of data)                              |
| Professional knowledge structure                                     | Stable and intuitive to user, but overlapping concepts can be a problem             |
| Type of data, photo, document, raw data, analyzed data, publications | Useful to some types e.g. photos, but often only used at sub category level         |
| Time   | Useful for long sequences such as breeding, else mainly e.g. as versioning of files |
| Geography  | Relatively stable, usefulness depends on importance of location                     |

#### Recommendation for folder

#### structure

- DEFINE A SYSTEM AND STICK TO IT
- IMEDIATELY SAVE NEW FILE IN PROPER PLACE IN STRUCTURE (NOT IN EMAIL, DESKTOP, MYDOC....)
- For each separately funded project, you want to keep everything related to that two folders (1 professional, 1 personal/contractual)
- Use the **knowledge structure of your professional work** on a high level as possible, and try to define non-overlapping stable categories
- Avoid if possible to use organizations and people as organizing principles as the tend to change e.g CIP is restructuring, but research topics stay the same

## Recommendation for folder structure

- Certain type of data such as photos, literature, GIS, you may want to have in special folder structured according to your professional logic
- Often your work has a sequence of improvement/processing and at a low level this should be reflected in the file name and possible in folders, e.g. use dates and PD (preliminary draft), 1D, 2D, FD for documents or raw data, cleaned data, analyzed data
- Consider using numbers to order your folders in logical sequence, rather than alphabetical

#### Example HD



12 CIP Jens

21 Litterature

22 Software

31 Privat

32 TIC MUNICIPAL

Downloads

InterVideo

🚮 My Music

My Music

My Pictures

My Pictures

My Videos

My Videos

Newsbin Download

Work and contract folder for CIP

Literature and software general professional categories across jobs

Private information and previous job

Folders required by different software installed and garbage from Windows

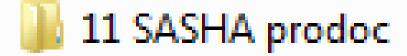
### Example HD - CIP SASHA folder

- 📗 1.1 Germplasm IS
- 2.1 Seed Systems IS
- 3.1 CSPro
- 📗 3.2 ME system
- 3.3 Nutritional studies
- 3.4 Qualitative data analysis
- 4.1 Collaborative platform
- 5.1 CIP data management
- 6.1 SASHA

1-4 specific areas of work

- 5. General area of work
- 6. Program/institutional work
- 7. Will probably add folder for other Programs or institutions

## Example HD - SASHA Program folder



21 Workplans

鷆 31 Budget

41 Reports

Typical content of Program/project folder

### File naming conventions

- Use **key words** that will help you in a search
  - Subject area, geographical region, organization
  - Type of document (Concept note, Budget, Agenda etc)
- Don't name a file by the person who will receive it or who gave it to you
- If certain **file types that are used repeatedly consider a formal naming convention** e.g. Series-Year-Season-Experiment (OFUG10A05)
- For non-standard files use longer names that indicates content and key words

### More file naming conventions

- Indicate version and date for docs with revisions
  - SP breeding Uganda Concept note Vo2 2010-05-10
- Consider how numbers and dates can order your files
  - 2010-05-24 Agenda vs 24 May 2010 Agenda
  - Trialo1, Trialo2, Trial10 vs Trial1, Trial10, Trial2
- NOTE: For formal Data Sets use the naming conventions of your institution and/or professional best practices

### Start reorganizing your harddisk

- 1. NEVER START RESTRUCTURING WITHOUT EXTERNAL COMPLETE BACKUP IN ORIGINAL STRUCTURE
- Create a folder "oldData"
- 3. Move everything into that folder
- 4. Note: If your harddisk is very full, you may have to work on only a part of the time

### Organizing the top level folders

- 1. Create new folder structure, TOP LEVEL
  - For each Project (funding source):
    - 1. Professional (information that will be shared)
    - 2. 1 Contractual/private (information not to be shared)
  - General folders
    - Literature folder
    - 2. Photos folder
    - 3. Software folder
    - 4. GIS data
    - 5. Video
    - 6. Private

#### Organizing Project folder

- For each Project Professional folder
  - 1 folder for each substantial area of work (non overlapping)
  - 1 folder for cross cutting issues in your work in the project
  - 1 folder for general project issues (prodoc, workplan, budget, meetings, etc)
  - Possibly one folder for collaborations with other projects, programs, institutions

#### Organizing Project folder - contract

- For each project (if relevant)
  - Folder for contractual issues
  - Folder for travel claims
  - Folder for time sheets

#### Other issues to consider

- Consider to use numbers to order folders logically
- Each folder should have some 3-10 subfolders
  - Less than 3 subfolders consider if you need main folder (reduce number of clicks)
  - More than 10 consider to group to avoid scrolling (scrolling is slower than clicking!)