A conceptual model for extraction of relevant geospatial data in *ChangeHabitats2* project

Balázs Székely[1], **Adam Kania**[2], **Norbert Pfeifer**[1]

balazs.szekely@ipf.tuwien.ac.at

with contribution of **János Tamás**[3], **Nikolett Szöllősi**[3], **Wojciech Wahlig**[1] and **Karolina Królikowska**[4]

[1] Institute of Photogrammetry and Remote Sensing (IPF), Vienna University of Technology, Vienna, Austria

[2] Atmoterm S.A., Opole, Poland

[3] Department of Water and Environmental Management, University of Debrecen, Hungary

[4] formerly at Atmoterm S.A., Opole, Poland

Rationale

"*Conceptual model for extraction of relevant geospatial data in CH2*"

conceptual: may apply to all activities without limiting them

extraction : procedural steps that create N2K relevant information

+ relevant : from remotely sensed data

geospatial: the extracted data are readily *integrated* in GIS

- Why at all?
  - Remote Sensing + Natura 2000 – powerful combination, but multitude of activities within *ChangeHabitats 2 (CH2)* cover a number of disciplines, somewhat puzzled
  - Habitat change: evaluation of field data + auxiliary information ➔ decision
  - Integrative approach needed, which does not violate any specific areas
- Why now?
  - Some conceptual/strategic decisions should also be taken in advance
  - Since many secondments are running currently, results will be available soon
  - It may provide a useful framework for all activities, enabling co-operation

Rationale

- What?
  - Create a common frame that defines some interface between subprojects
  - Categorization of activities and recognizing missing links, research needs
  - These documents will be integrated into a larger conceptual framework
- Expected results:
  - Common understanding, applicable internal standards
  - Identification of not yet recognized derivatives of remotely sensed data
  - Seeds of scientific paper(s)
  - Precursor of how-to-s to be written for specialists involved in Natura 2000
  - It will also help report creation + co-operation with stakeholders

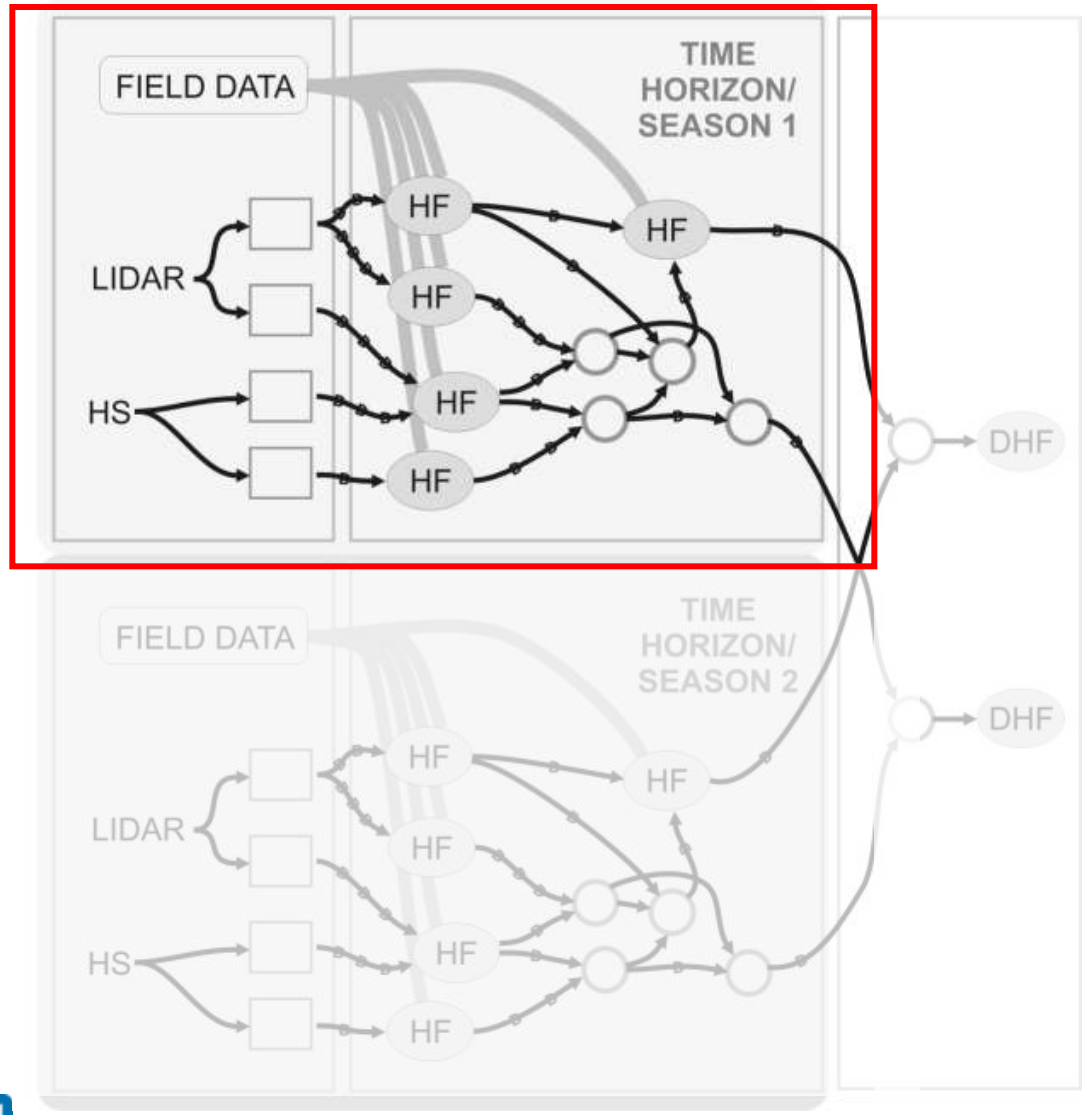
Definitions

- Habitat Feature (HF): *any feature of the N2K site that can be observed, measured, calculated or implied (regardless if it is a value, a judgement or a state), that is relevant for the evaluation of the habitat in the sense of the 1992 Habitat Directive and related EU or national legislation*
  - Explicit HF: if the quality/value/state is to be entered into the data form of the N2K site *example:percentage of a habitat in a N2K site*
  - Implicit HF: if the quality/value/state/observation may be used to calculate or imply one or more explicit HF
    example:amount of dead wood e.g. in m³
    (assuming it is important at the actual N2K site)
- Change of a HF (DHF): *temporal variation of an explicit HF*, especially important if this change indicates deterioration of the N2K site

HFs are determined
partly by definition of the legislation, and
partly by the research and practice that describe how to calculate it

The conceptual framework

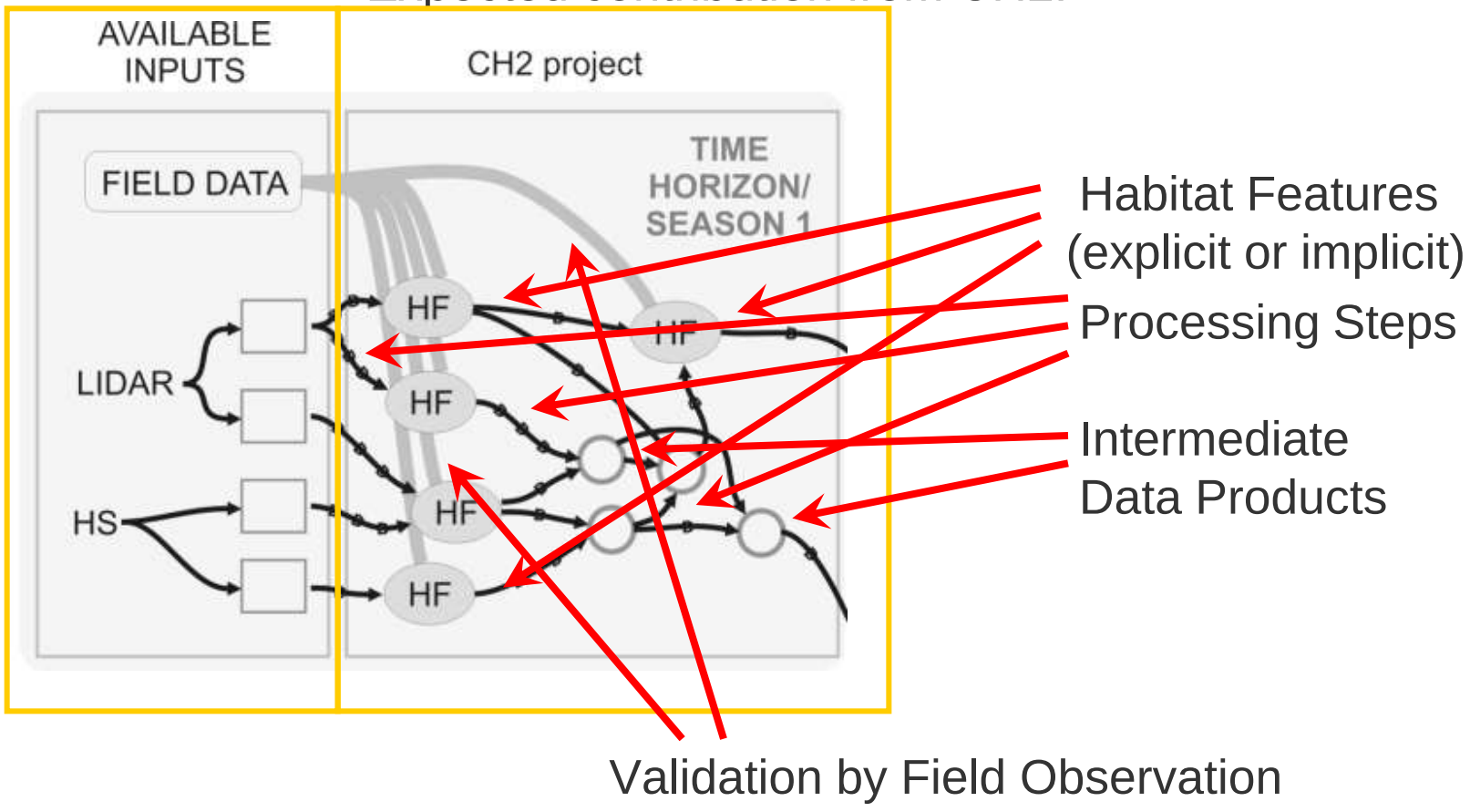
AVAILABLE INPUTS | CH2 project | CH2 results



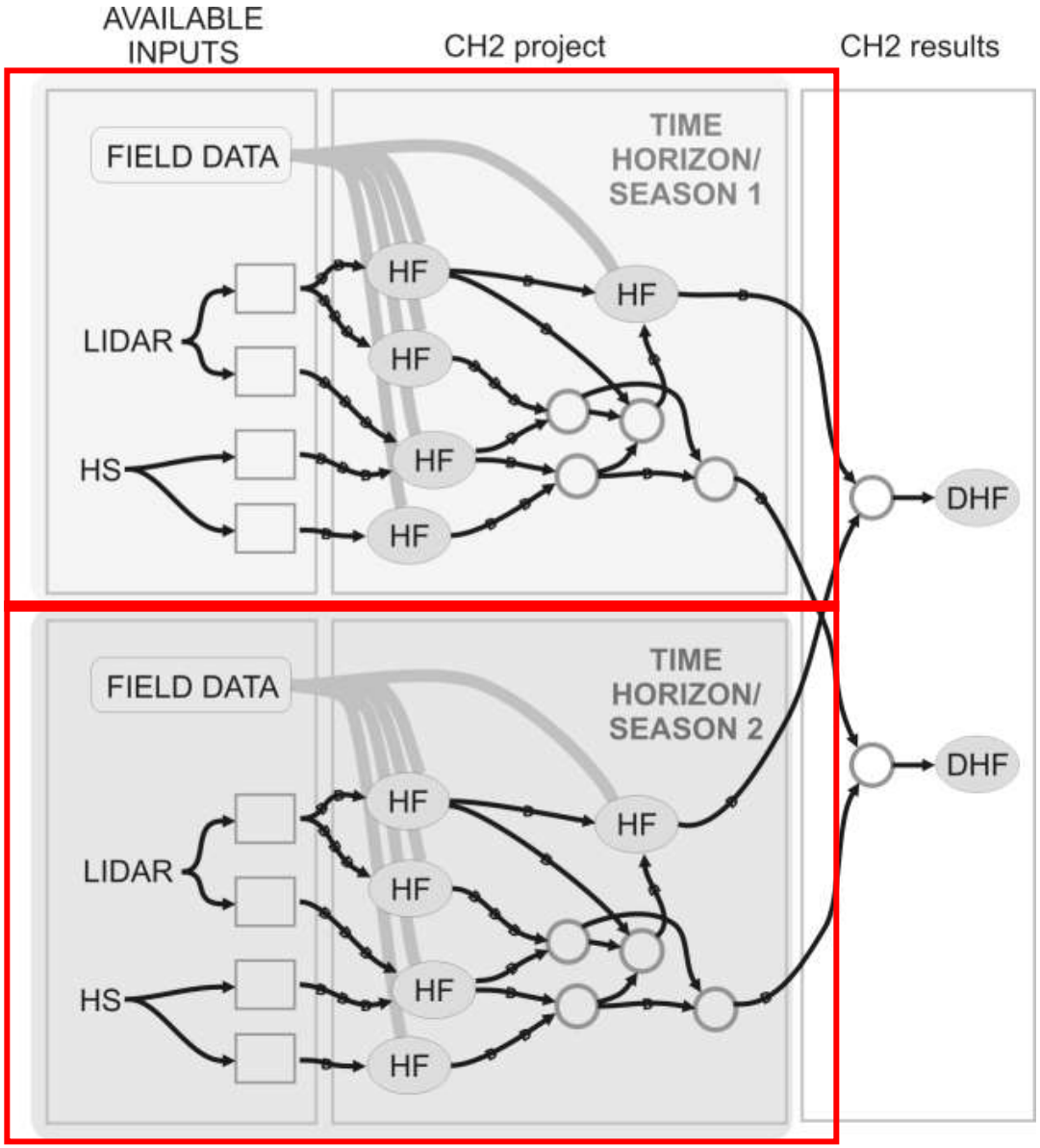
The conceptual framework: single timeslice

That exists today:

Expected contribution from CH2:



The conceptual framework: multitemporal assessment



Multitemporal:

optimal, if the sequence of observations happen under the same circumstances

but

the concept should allow that the observations are made under different conditions (instrument, operator, resolution, etc.)

The conceptual framework: multitemporal assessment

Isn't it an erroneous design?

It is possible that differing circumstances may hamper a simple (direct) comparison (e.g. change detection) of the observations, because they could be *per se* incomparable due to the differing conditions.

However, the solutions are intended to be applied at typical N2K sites, where not all circumstances are favorable.

The comparability of the observations in these cases is achieved by using corrections, calibrations, modelling or similar techniques, if possible.

This also means that the interpretative level of various (data) products will be different.

Multitemporal:

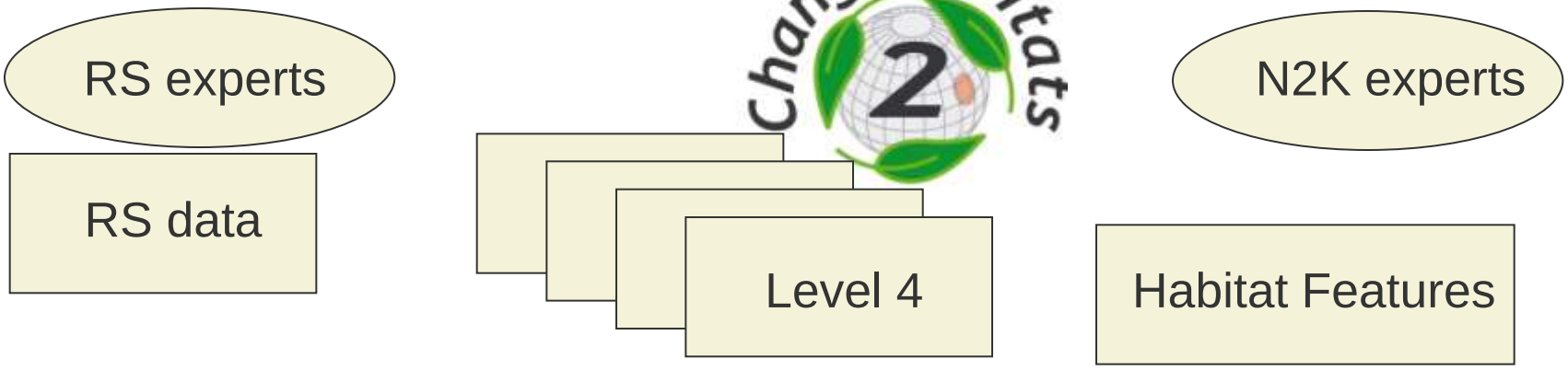
optimal, if the sequence of observations happen under the same circumstances

but

the concept should allow that the observations are made under different conditions (instrument, operator, resolution, etc.)

Data Processing Level

Necessary to introduce the data processing levels, similarly to other RS applications.



Processing Levels

Levels of Data Processing

NASA's Definition for Levels of Data Processing

- **Level 0** Reconstructed, unprocessed instrument/payload data at full resolution; any and all communications **artifacts**, e.g., synchronization frames, communications headers, duplicate data **removed**.
- **Level 1A** Reconstructed, unprocessed instrument data at full resolution, time-referenced, and **annotated with ancillary information**, including radiometric and geometric calibration coefficients and georeferencing parameters, e.g., platform ephemeris, computed and appended but not applied to the Level 0 data.
- **Level 1B** Level 1A data that have been **processed to sensor units** (not all instruments have Level 1B data products).
- **Level 2 Derived** geophysical **variables** at the same resolution and location as the Level 1 source data.
- **Level 3** Variables **mapped on uniform space-time grids**, usually with some completeness and consistency.
- **Level 4 Model output or results from analyses** of lower level data, e.g., variables derived from multiple measurements

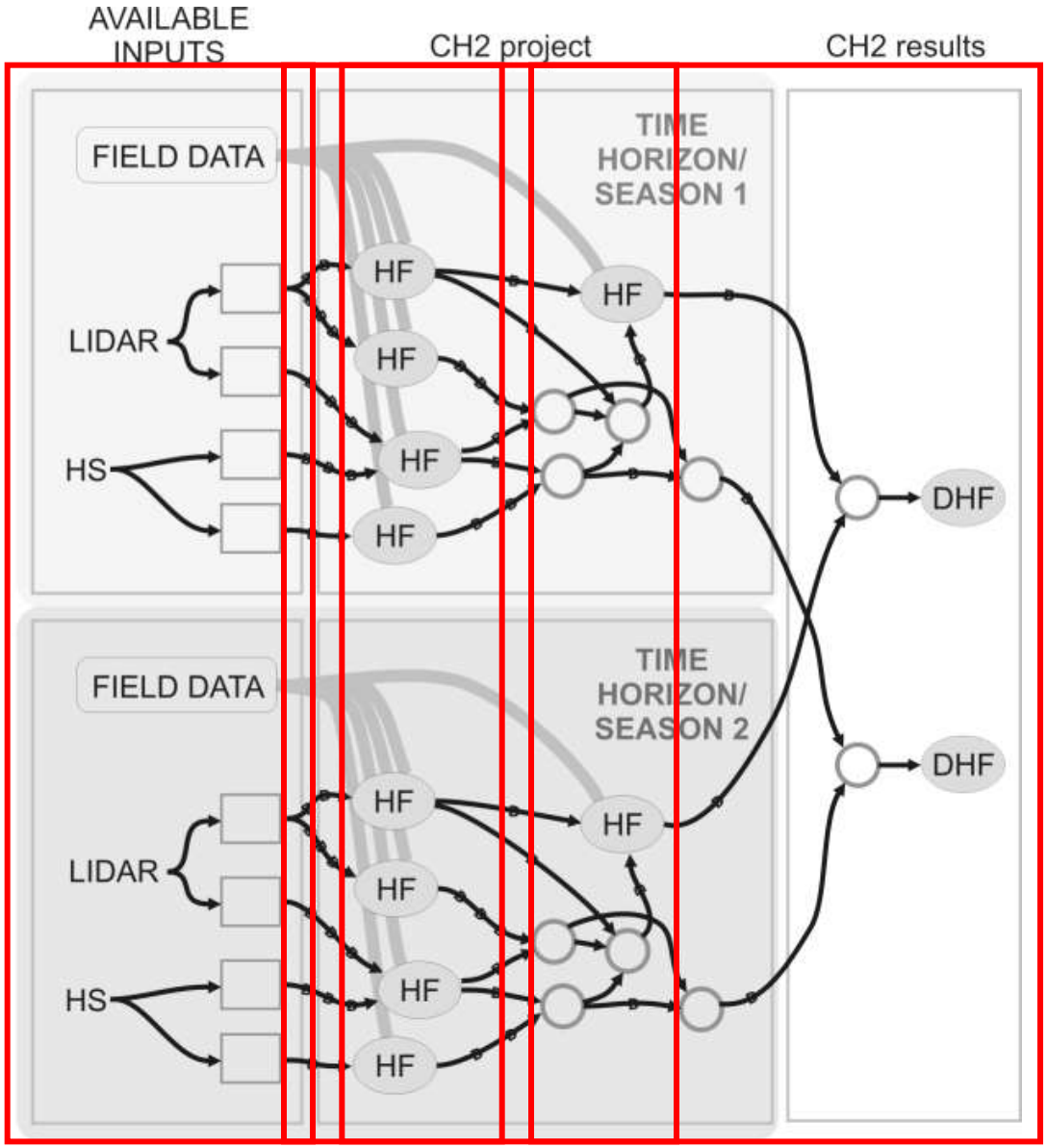
Parkinson & al. (Eds.), 2006: Earth Science Reference Handbook – A Guide to NASA's Earth Science Program and Earth Observing Satellite Missions

Proposal for Processing Levels in CH2

Level 0 – Level 1B	industry-standard data processing steps, yet not influenced by the habitat-research-specific requirements
Level 2 **(iHF)**	products are those derivatives that can be still produced by industry-standard methods or by methods published in the scientific/technical literature, but these methods were selected and tailored by the habitat research requirements
Level 3 **iHF**	products are data constructs that were generated by using specifically designed and implemented processing tools, concepts, evaluation steps and decision trees that were developed for the habitat research
Level 4 **eHF**	constructs are those objects that can be directly used in N2K habitat change evaluation as it is driven by Habitats Directive

iHF:implicite HF **eHF**:explicit HF

The conceptual framework: multitemporal assessment



Level 0 – Level 1B

Level 2

Level 3

Level 4

Conclusions

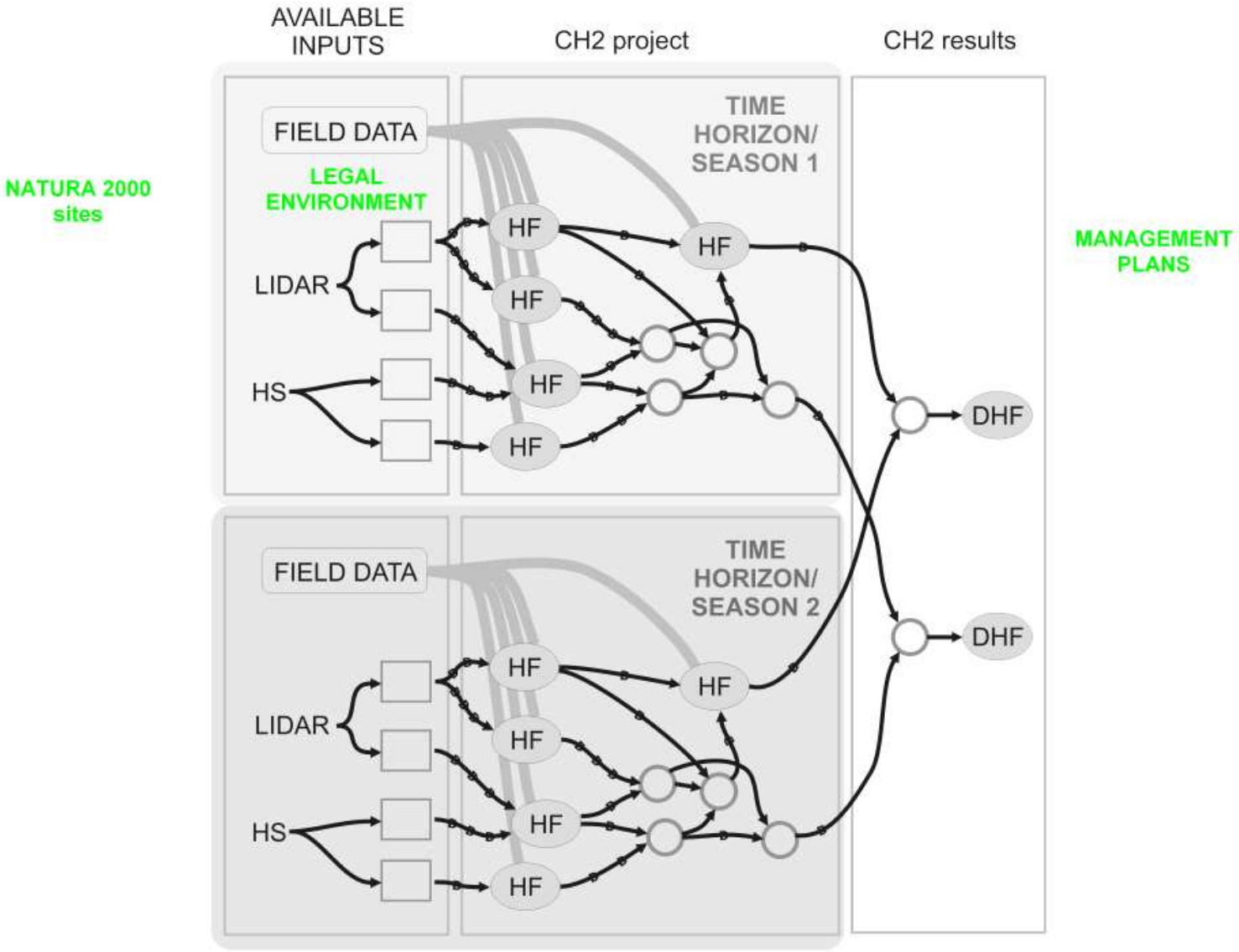
- All CH2 activities can be inserted into the scheme
- The scheme allows straightforward visualisation and communication of the subprojects
- Processing chains and dependencies become obvious for all participants
- Research needs can be detected
- The role of the activities can be related to the Habitats Directive, this helps in convincing potentially skeptic users/N2K managing entities

Thank you for your kind attention.

balazs.szekely@ipf.tuwien.ac.at

BS wishes to express his gratitude to the colleagues and their family members who have removed all bureaucratic and practical obstacles and made his two-month-long secondment in Poland a fruitful stay and an unforgettable memory.

This slide is intentionally empty.


AVAILABLE INPUTS
CH2 project
CH2 results
NATURA 2000 sites
FIELD DATA
LEGAL ENVIRONMENT
LIDAR
HS
TIME HORIZON/ SEASON 1
HF
HF
HF
HF
HF
DHF
MANAGEMENT PLANS
FIELD DATA
LIDAR
HS
TIME HORIZON/ SEASON 2
HF
HF
HF
HF
HF
DHF

Selected terms 1

- geospatial data
  - two lanes: (1) sense of CH2 questionnaire (2) internal sense of CH2
  - international definition attempts may also be used partly
- generalisation
  - role of generalisation – strong dependency on the meaning of scale
- raw data/derivative data
  - from technical point of view, derivative data are produced at the low levels (e.g. FWF LS: points are already derivative values)
  - from ecological point of view, such low level constructs are still considered to be raw data, because of the low ecological model content
- data/feature extraction
  - strongly linked to previous term
  - in a sense it is the real goal of the project: to figure out what and how
  - high model influence/model content

Selected terms 2

- data integration ([one,many] ←→ [one,many])
  - what is the entity? depends on in which sense, according to what method
  - how to handle if an entity from this point of view is split/fuzzy/structured while from another point of view (e.g., observational, biological, logical, legal point of view) it is different?
- visualisation, mapping
  - what do we (CH2) and what a (CH2-related|public) user considers visualisation? how far it is needed to be dynamic|static|pre-defined?
  - mapping: scale dependency, accuracy, data integration
- (multi)temporal data (what is multitemporal?)
  - conflicting ideas: is this multitemporal: e.g. LiDAR flights
    - on 30th May and on 10th Jun?
    - on 30th May and on 2nd Aug?
    - on 2nd Jan and on 30th May?
    - on 11th Nov, on 4th Feb and 5th Jun?
  - and the same dates for HS flights? (consider also sun-target-sensor angle)

Selected terms 3

- handling time, change
  - concept of events
    - Is a forest fire/landslide a development step? How to distinguish?
  - the time passed between RS acquisitions and between RS and field survey
  - how far the time passed between RS acquisitions can be used for extrapolation/interpolation?
- qualitative states/quantitative properties (also compare to scale!)
  - instead of quantitative properties sometimes only qualitative states/categories/classification can be given
- comparison of multitemporal data: data/position accuracy (differing/same)
  - distinction of “change” because of differences in data acquisition resolution (depending on e.g. flight height), accuracy, positional accuracy

Selected terms 4

- projection issues
  - scale dependent: UTM is good for local representation
  - countrywide representation needs e.g. Lambert Azimuthal Equal Area proj.
- 1D,2D,2.5D,3D,(4D)
  - 1D, 2D (profiling, mapping)
  - 2.5D the mapped volume (e.g. laser scanning) is very thin
  - 3D only on local, microscale
  - (4D) including multitemporality, but maybe “just” mapping
- sampling theorem, fractal structure, spacing, data scatter
  - $f_s > 2\, f_{Nyquist}$
  - fractal behaviour on several order of magnitudes, pattern
  - spacing (cf. fractal behaviour)
  - data scatter (cf. sampling theorem)
    - how far the data acquisition is evenly distributed?
    - is there any implication on the comparison and the explanatory power?

Selected terms 5

- feature: existence/presence/percentage/ubiquity
  - existence: can we detect a one single feature
  - presence: low abundance, but detectable (what if not)
  - percentage: estimation of percentage and its error
  - ubiquity: what is the threshold for that?
- related intervals/boundaries: fuzzy, sharp, defined
  - fuzzy boundary
    - because of the phenomenon
    - because of the data acquisition
  - sharp boundary
    - what accuracy is required?
  - defined boundary
    - can we verify boundaries that are by definition boundaries?

Rationale

- Why at all?
  - Activities within *ChangeHabitats 2 (CH2)* cover a number of disciplines
  - The usage of terms in various disciplines is/may be different
- Why now?
  - Some conceptual/strategic decisions should also be taken in advance
  - Since many secondments are running currently, results will be available soon
- What?
  - Create a list of some fundamental terms that influence all/many WPs
  - Start discussion about these terms by formulating short documents
  - These documents will be integrated into a larger conceptual framework
- Expected results:
  - Common understanding, applicable internal standards
  - Seeds of scientific paper(s)
  - Precursor of how-to-s to be written for specialists involved in Natura 2000

Suggested proceeding

- How?
  - Some tools will be available soon on the CH2 server
    - SVN – data storage/exchange facility
    - Interactive forum – may replace many cross-posted e-mails
    - Web info (separated info also for internal use – TBD what) dynamic development
  - Proposal: introduction of CH2–RFC* (*CH2 Request for Comments*) procedure
  - Creation of topical groups (Is it necessary? Are there volunteering people?)

* Note that "*CH2*–" is intentionally added to the name in order to distinguish our procedure from the internet RFC that somewhat differs.

Terms and Definitions

- ***application profile**: identification of clauses, classes, subsets, options and parameters from base standards that are necessary for accomplishing a particular function*
- ***client**: software component that can invoke an operation from a server*
- ***dataset**: identifiable collection of data [ISO19115]*
- ***dataset series**: collection of datasets sharing the same product specification [ISO19115]*
- ***geographic dataset**: dataset with features depicted geometrically*
- ***geographic information**: information concerning phenomena implicitly or explicitly associated with a location relative to the Earth [ISO19101]*
- ***identifier**: linguistically independent sequence of characters capable of uniquely and permanently identifying that with which it is associated [ISO19135]*
- ***metadata entity**: set of metadata elements describing the same aspect of data [ISO19115]*
- ***metadata schema**: conceptual schema describing metadata [ISO19101]*
- ***metadata section**: subset of metadata that defines a collection of related metadata entities and elements [ISO19115]*
- ***operation**: specification of a transformation or query that an object may be called to execute [ISO19119]*
- ***parameter**: variable use to express a value in an operation request or response*
- ***request**: invocation of an operation by a client [ISO19128]*
- ***response**: result of an operation returned from a server to a client [ISO19132]*
- ***schema**: formal description of a model [ISO19101]*
- ***server**: a particular instance of a service [ISO19128]*
- ***service**: distinct part of the functionality that is provided by an entity through interfaces [ISO19119]*

NAP-MWG, NAP – Metadata Working Group (2007): North American Profile of ISO19115:2003 - Geographic information – Metadata (NAP – Metadata, version 1.0.1), http://www.fgdc.gov/standards/projects/incits-l1-standards-projects/NAP-Metadata/napMetadataProfileV101.pdf/view, accessed on 13/07/2011.

Concept of János Tamás