

# Transparent and reproducible processing of model inputs with flexible spatial aggregation

Lavinia Baumstark, Jan Philipp Dietrich, Anastasis Giannousakis, David Klein

PIK RD III Sustainable Solutions

## CHALLENGES OF INPUT DATA PREPARATION

In modelling, the preparation of input data is often one of the challenges. Model developers are often faced with...

- ... raw data in many different structures, formats and resolutions
- ... unstructured and hardly reproducible input data preparation
- ... different programming styles in a team of developers

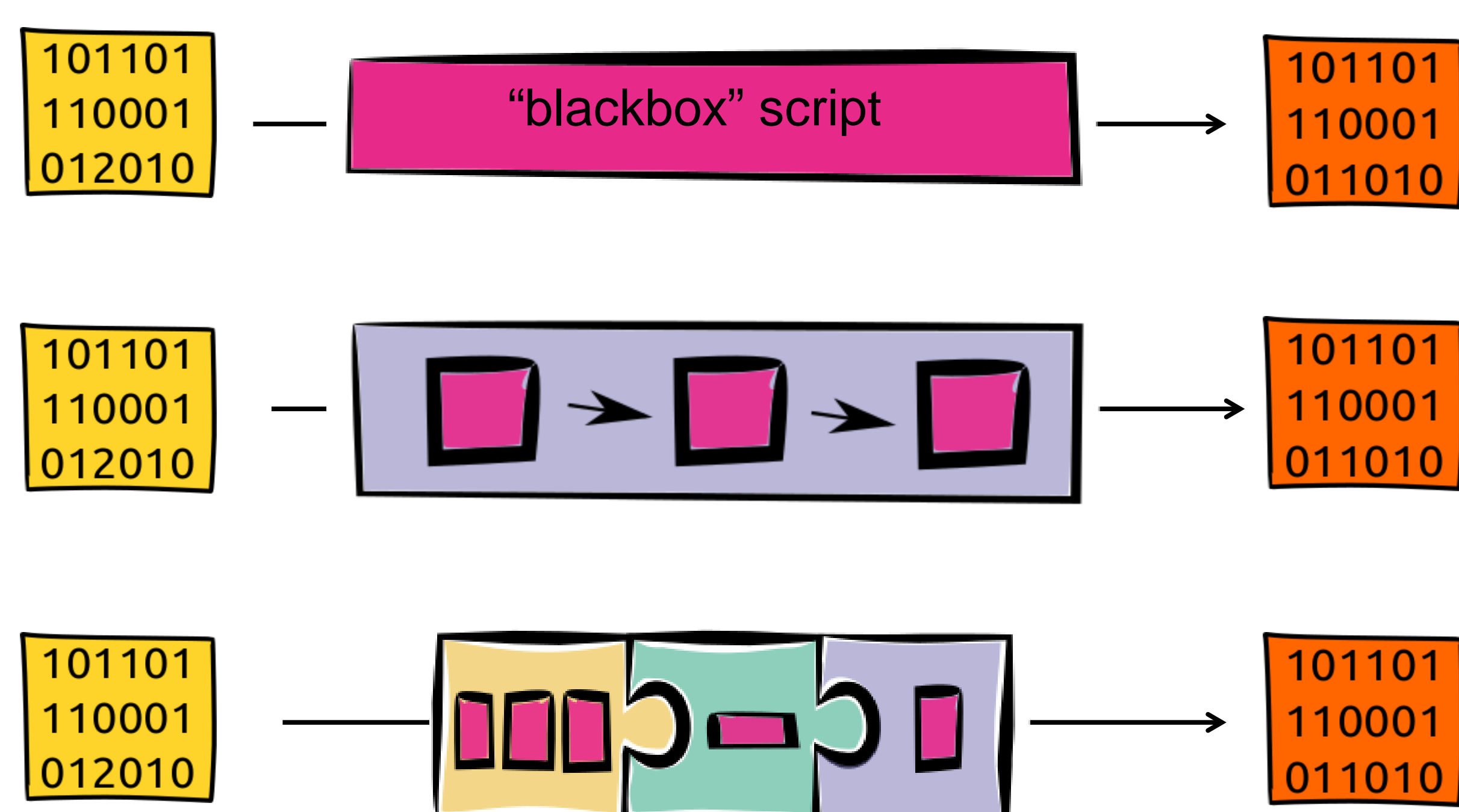
## R-Package “MADRaT”

“May All Data be Reproducible and Transparent”

Reference: Dietrich J.P., Baumstark L. and Giannousakis A. (2018).  
\_madrat: May All Data be Reproducible and Transparent (MADRaT)\_  
Doi: 10.5281/zenodo.1115490 R package version 1.49.0

Git: <https://github.com/pik-piam/madrat>  
CRAN: <https://CRAN.R-project.org/package=madrat>

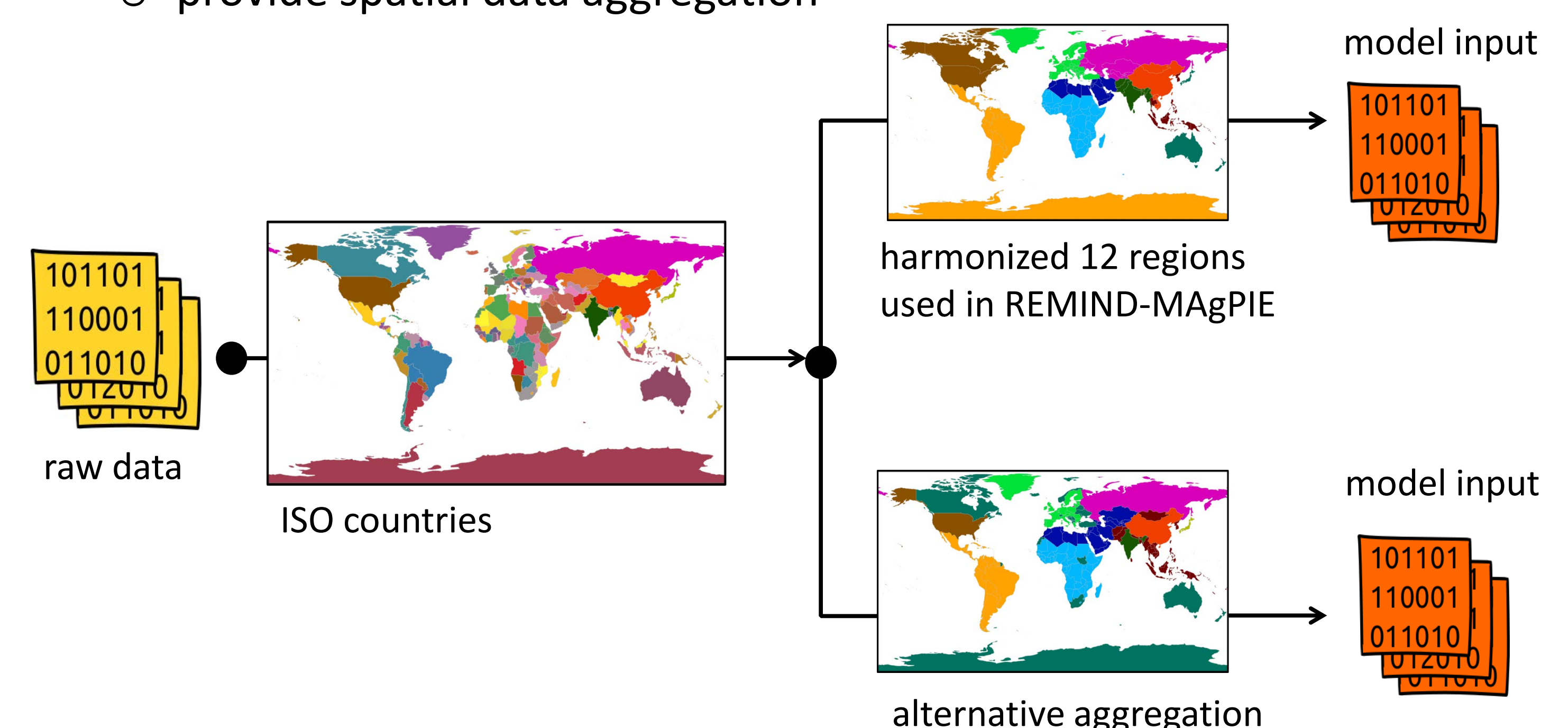
## INTRODUCING STRUCTURE VIA WRAPPER FUNCTIONS



## INTRODUCING SERVICES VIA WRAPPER FUNCTIONS




The wrapper functions cover various additional tasks of data processing:

- apply consistency checks on the data
- create and manage meta data
- handle data caching
- provide spatial data aggregation



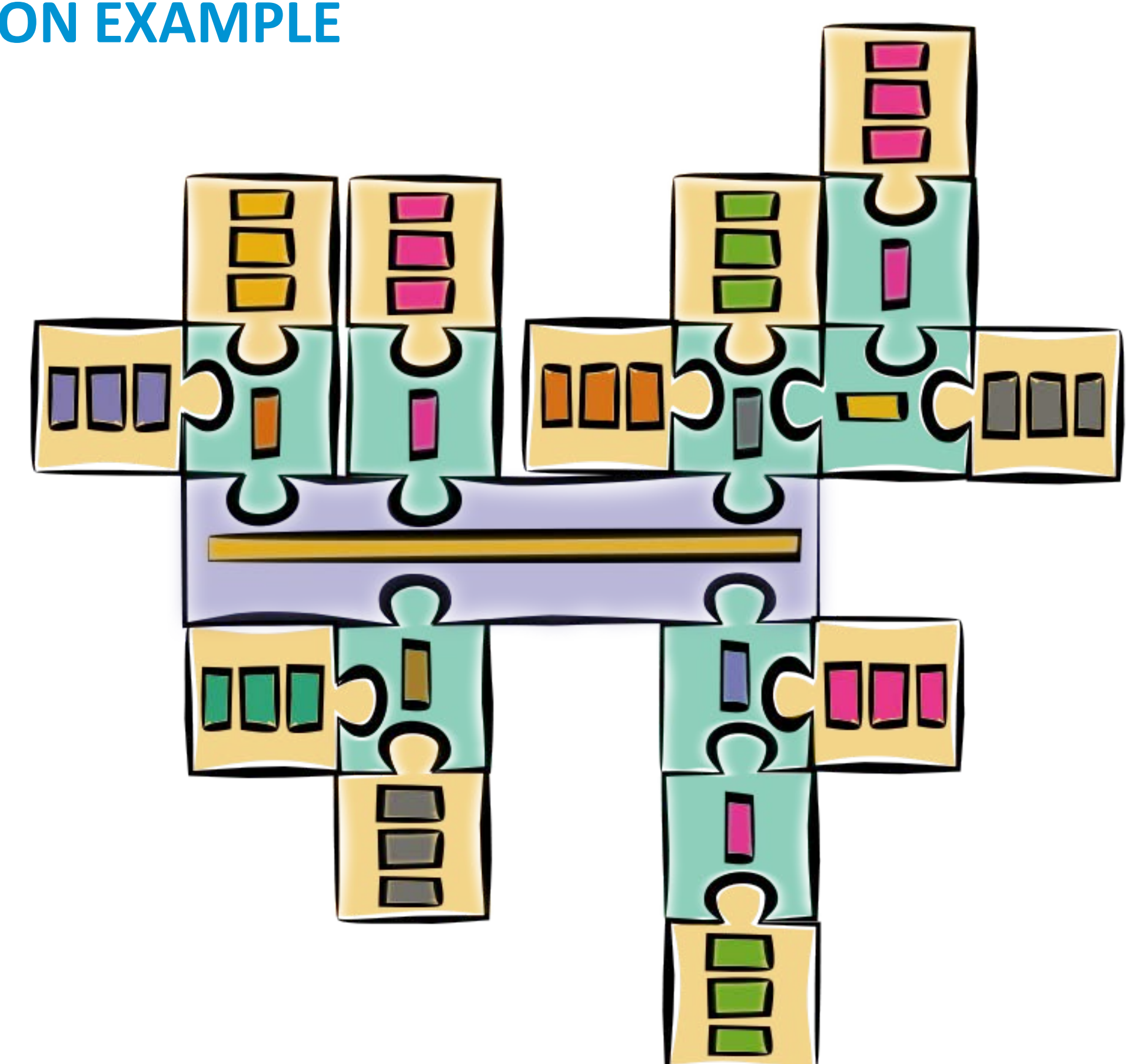
## PROCESSING STEPS

Input data preparation is split into three distinct processing steps with clear interfaces:

-  **readSource**
  - download source data
  - import and convert data to a standardized data format
  - bring data to the ISO country level
-  **calcOutput**
  - calculate required data
  - filter and merge data (from different data sources)
  - provide information for spatial aggregation
-  **retrieveData**
  - collect data sets
  - coordinate packing of aggregated data

For each of these tasks wrapper functions are provided.

## APPLICATION EXAMPLE



An application may consist of many different components which might be called more than once.