

**Dominik Lin**  
Economist-Statistician  
DGS

# **BIRD - a formal description of reporting requirements based on the input approach**

***DISCLAIMER:** The views and opinions expressed in this presentation are those of the author and do not necessarily represent official policy or position of the ECB.*

# Overview

## Introduction to the BIRD

- BIRD process

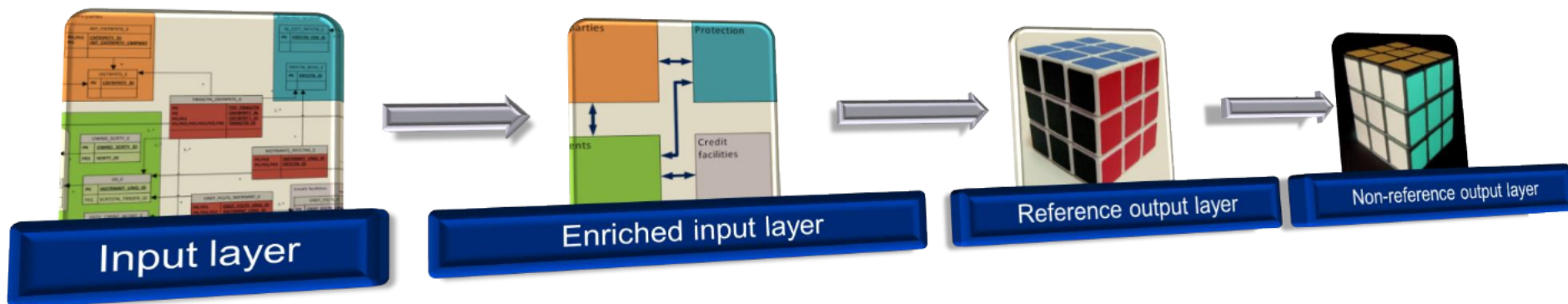
## Description of the dictionary content

- Cubes, Mappings & Transformations

## Merits & costs

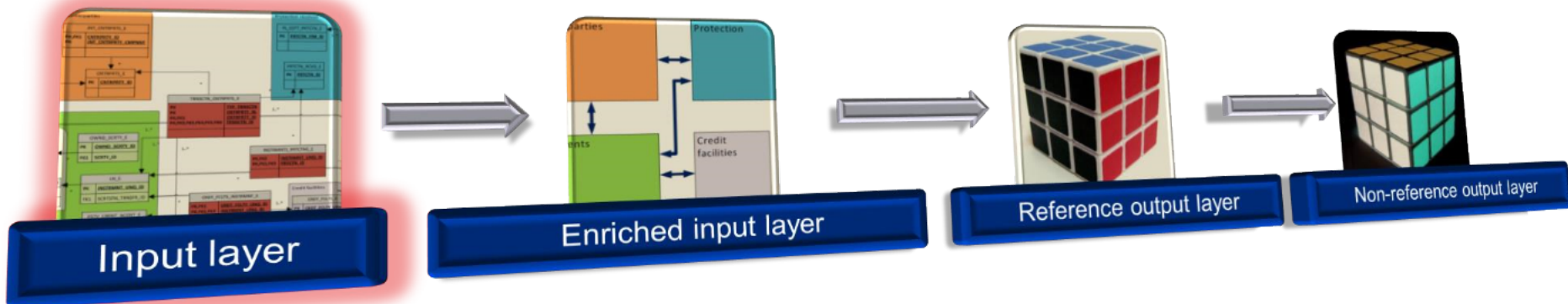
- Of a BIRD like approach

## Process overview



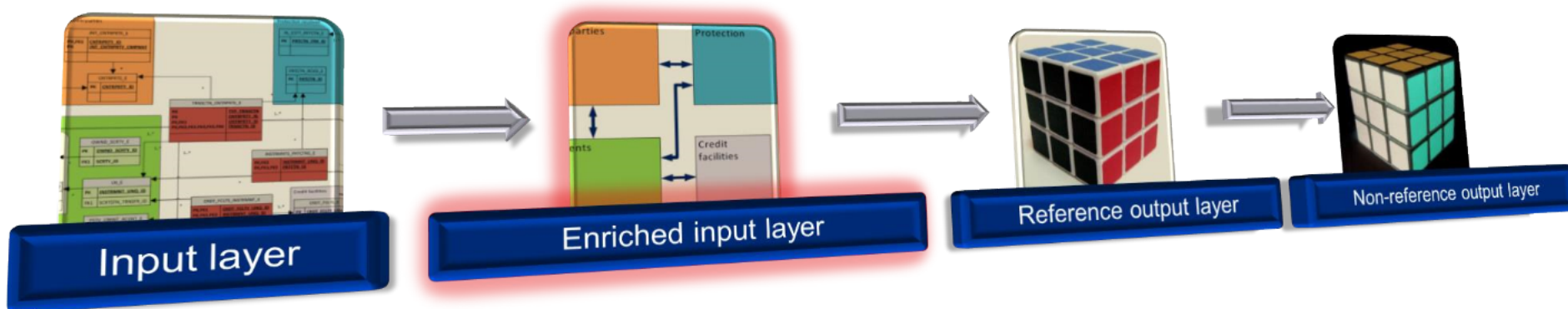
- Process to generate the output layer cube(s) by applying specific transformations onto the input layer (data)
- It's components are described using
  - Cubes
  - Transformations
  - Mappings

## The input layer



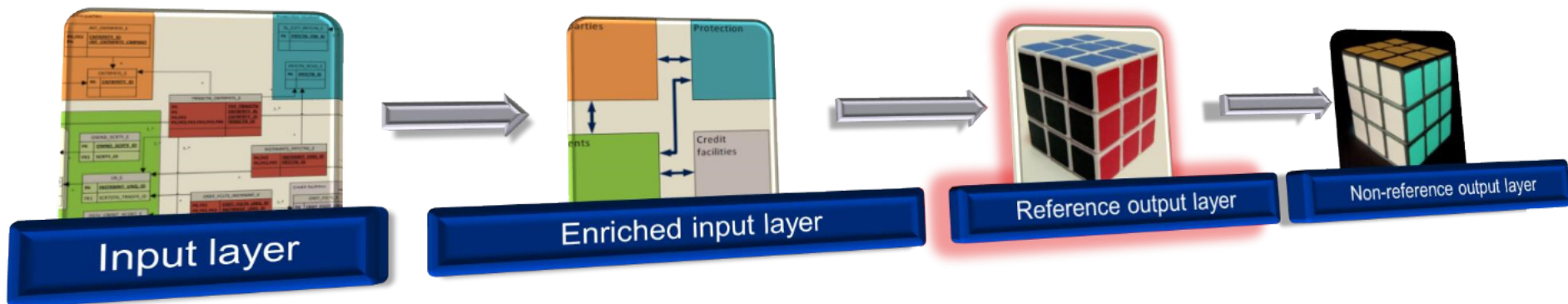
- Entity Relationship Model (ERM) comprising information that is relevant for different reporting purposes (i.e. output layer)
- Interface to bank's internal systems
- One model (i.e. input layer) in order to generate different reports (i.e. output layers)

## The enriched input layer



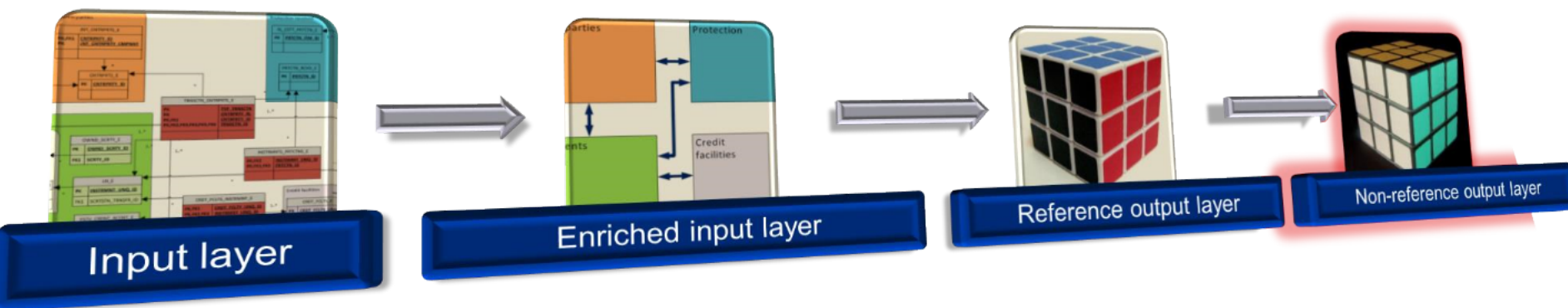
- Intermediate layer between the input and the output layer(s)
- Enriched information (e.g. result of the derivation of the SME classification)

## The reference output layer



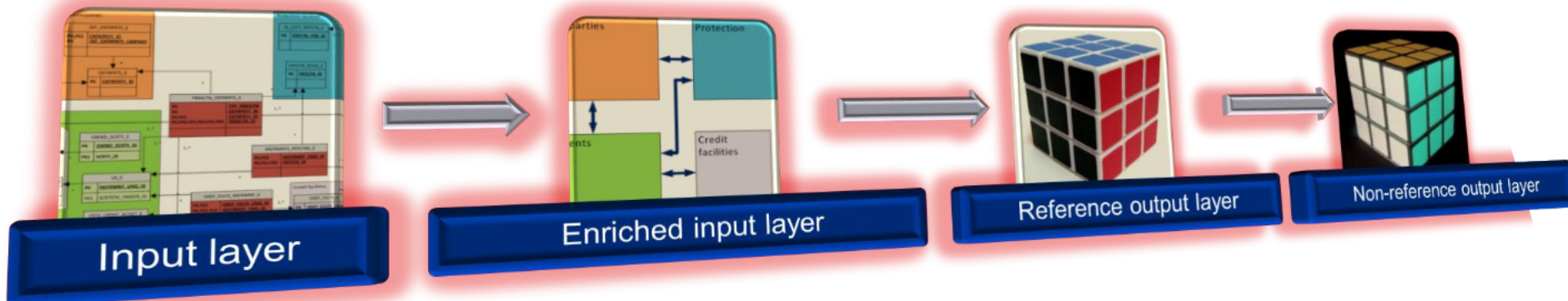
- Description of the output layer using *reference codes* (i.e. the codes used in the input layer)

## The non-reference output layer



- Description of the reporting requirements in the dictionary (i.e. using data sets)
- May be defined using a different codification system (i.e. non-reference) than the one that is used in the input layer (i.e. reference)

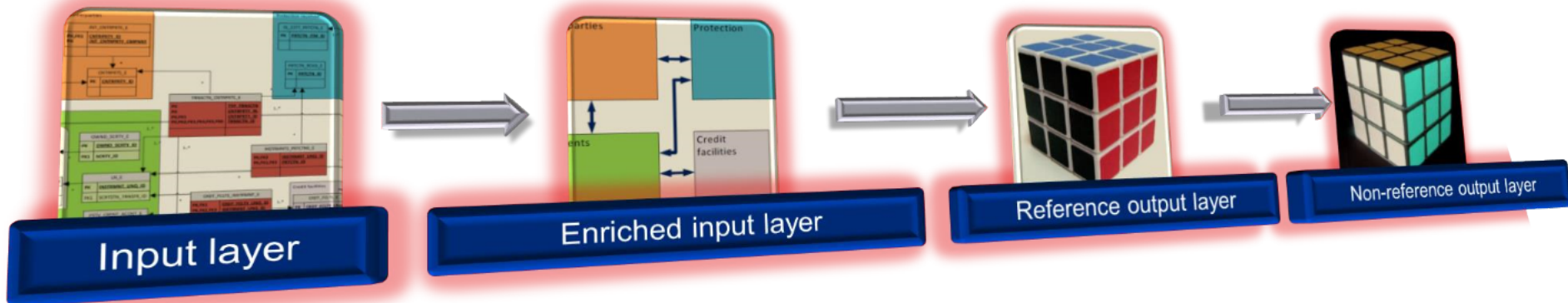
## Dictionary



- Based on SMCube methodology
- Provides functionalities
  - Reusability of concepts / univocal identification of concepts
  - Historization / versioning
  - Extensibility
  - Different codification systems and Mappings translating between those codification systems
  - Compatible with SDMX and XBRL

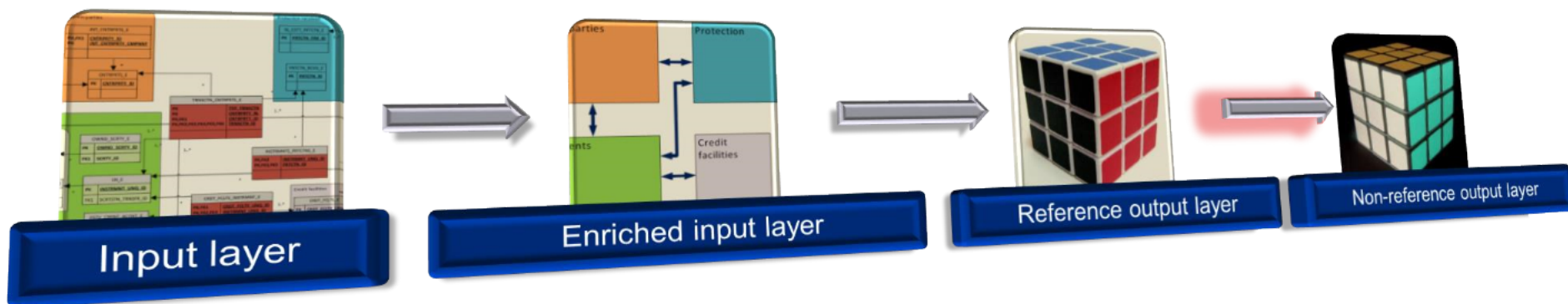


## Cubes / data definition



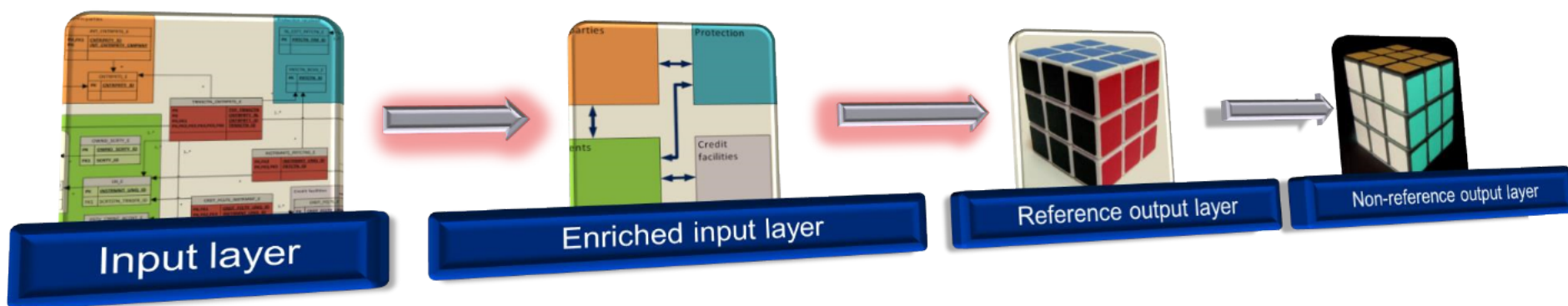
- Description of data sets (based on their Meta data)
- Comprises all important information about the data set (e.g. to generate the related data base table)
  - Variables (or columns) and their allowed values
  - Composition of the primary key
  - (one-to-many) relationships to other data sets via foreign key constraints
- Cube versions may be consumed by Meta data driven systems

## Mappings



- Mappings refers to the relation between the reference and the non-reference output layer
- Could be applied for the interface between a bank's internal system and the BIRD input layer in case of different codification systems

## Transformations



- Describes the necessary operations in order to transform one data set into another data set
- Univocal representation of each operation by
  - Applying the SDMX information model for transformations
  - Definition of the tree structure of each operator
- A set of transformations (i.e. Transformation Scheme) is describes using
  - Natural language (business description)
  - Validation & Transformation Language (VTL)

## Overview

- **Integrated** due to the input approach and reusability of concepts
  - Consistent reports (one input layer generating multiple output layers)
- Managing **data flows** – graph analysis of cubes & transformations
  - Parallelization of subgraphs
  - Impact analysis on cube / variable level in case of changes in reporting regulation
    - Which input cubes contribute to a specific output cube?
    - To which output cubes contributes a specific input cube?
    - Which input variables contribute to a specific output variable?
    - Which variables need to be reported at which reference dates?
    - Data quality optimization
- Sharing the analysis of the regulation, manuals, guidelines
  - Unique formal description for the reporting agents
- Thin line between documentation and implementation (**logical vs. technical model**)

### Input approach

- **Integrated** due to the input approach and reusability of concepts
  - Consistent reports (one input layer generating multiple output layers)
  
- The dictionary needs to be managed properly
  - Keep the dictionary clean
    - In case of new requirements the person responsible for the maintenance of the dictionary need to ensure that the dictionary is not corrupted (e.g. duplication of concepts)
    - Dependencies
  - Specific skill set for dictionary administrators
    - Know how about the underlying information model (of the dictionary)
    - Business knowledge
  
- Output requirements need to be represented using the same methodology / information model

# Formality of cubes & transformations

- Managing **data flows** – graph analysis of cubes & transformations
  - Parallelization of subgraphs
  - Impact analysis on cube / variable level
    - Which input cubes contribute to a specific output cube?
    - To which output cubes contributes a specific input cube?
    - Which input variables contribute to a specific output variable?
    - Which variables need to be reported at which reference dates?
    - Data quality optimization
- The dictionary needs to be managed properly
  - Specific skill set for dictionary administrators
    - Logical understanding of data manipulation in order to formulate transformations in a formal way

### Level of formality (1)

- Sharing the analysis of the regulation, manuals, guidelines
  - Unique formal description for the reporting agents
- Is done by each bank individually, a BIRD like solution would at least harmonize the solution

### Level of formality (2)

- Thin line between documentation and implementation (**logical vs. technical model**)