



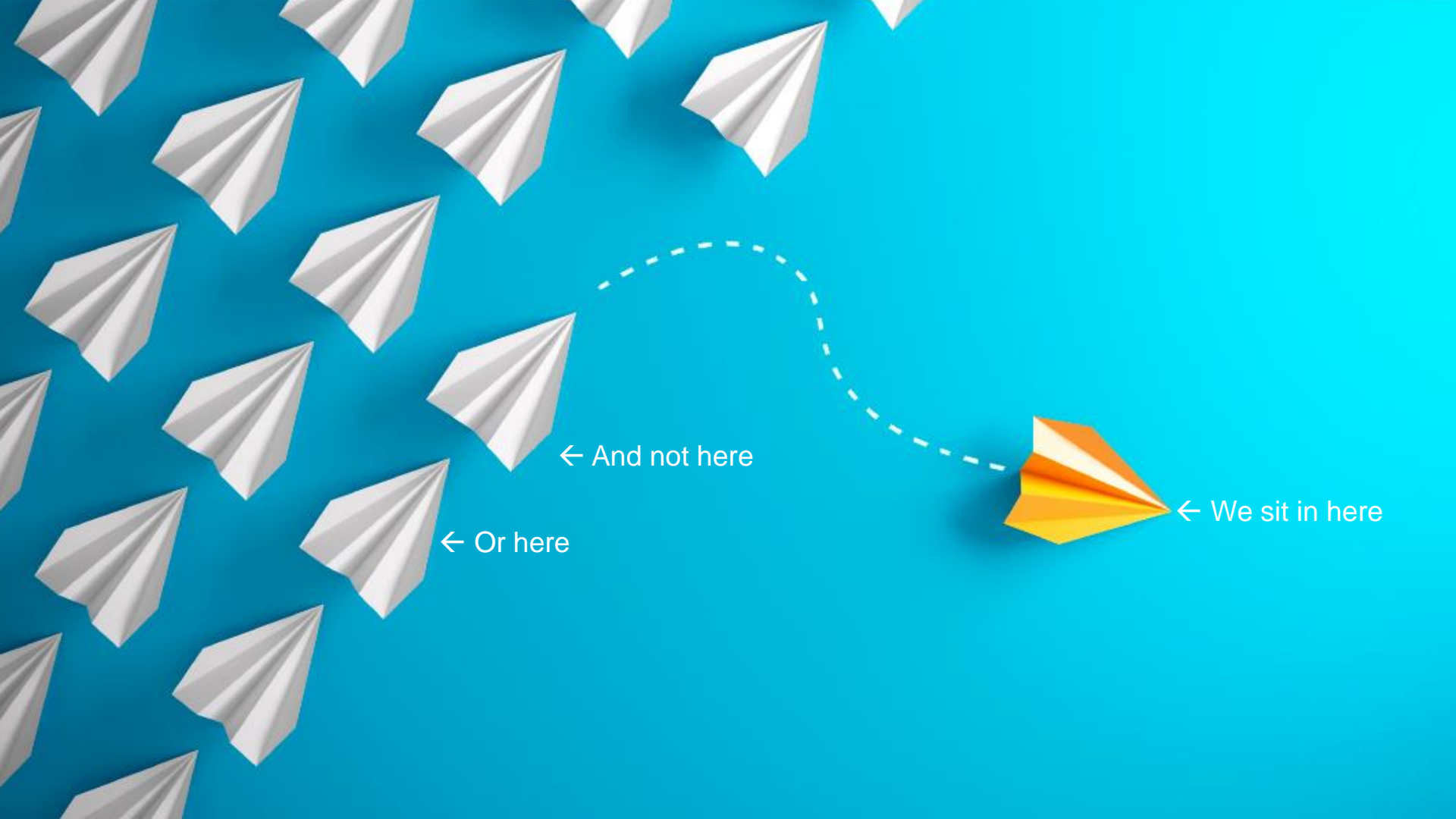
Metadatamodels like DPM

Why it might be a building block regarding a potential AI strategy

#BelieveInYourself



Christian Ronacher
Christian.ronacher@erstegroup.com

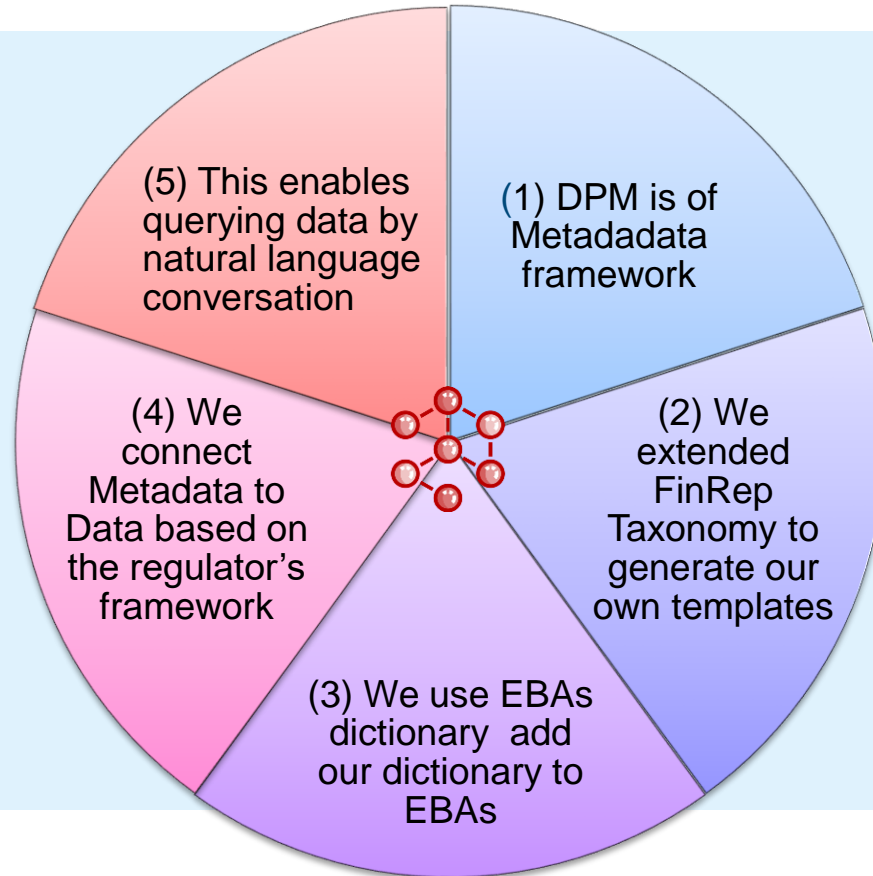


← And not here

← Or here

← We sit in here

What we do with DPM



Our goal is to provide
self explainable and
curated data,
which can be queried by
conversation.

To achieve this,
data must meet
the following criteria.

Consistent

Transparent

Efficient

Validated

Simple
accessible

We think metadatamodels
are a cornerstone
to provide meaning to reports
and make them queryable **by conversation**
(in an auditable way)

Sidestep – What do we mean with querying by conversation

GPT 4.0

The problem

C what is the carrying amount?



The carrying amount is not explicitly provided in the text you shared. However, the carrying amount for these derivatives can be calculated by taking the difference between the positive fair value and the negative fair value.

For interest rate risk derivatives held for trading in the trading book, the carrying amount as of December 2022 would be:

Carrying amount = Positive fair value - Negative fair value

Carrying amount = 5,788 million EUR - 5,508 million EUR

Carrying amount = 280 million EUR

Please note that this is a simplified calculation, and there may be other factors to consider in a more comprehensive approach.

This needs to be trackable

„reports follow the question“ vs. „I have know which report answers a certain question“

AL

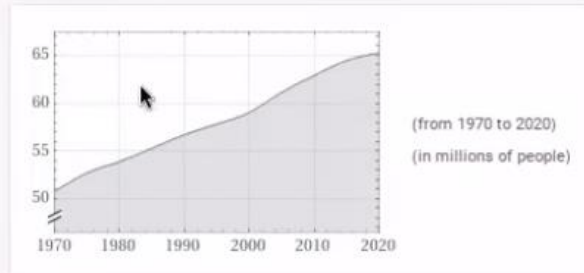
What's the population of france?



Used Wolfram



As of the 2020 estimate, the population of France is approximately 65.3 million people, making it the 22nd most populous country in the world. Here is a chart showing the recent population history of France:



The population density of France is estimated to be 309 people per square mile. The population growth rate is approximately 0.14% per year, and the life expectancy is 82.8 years. The median age of the population is 41.6 years.

Source: <https://youtu.be/EOQV9VakBgE>

Where do we come from?

Template

F 01.01		Carrying amount
Balance Sheet Statement [Statement of Financial Position]: Assets		010
010	Cash, cash balances at central banks and other demand deposits	10898
020	Cash on hand	112718
030	Cash balances at central banks	151111
040	Other demand deposits	151115
050	Financial assets held for trading	11110
060	Derivatives	11108
070	Equity instruments	11112
080	Debt securities	11106
090	Loans and advances	11114
091	Trading financial assets	45921
092	Derivatives	45920

Map each cell

	Carrying amount
	010
	10898

Documentation:
in this cell the carrying amount for cash and cash equivalent positions

Write Code for the cell

Cell 10898:

```
SELECT *
FROM xyz Table
Inner Join ....
```

Documentation:

This select statement retrieves BET_100 (carrying amount) for WER_ART_103 („Cash“ or „Cash equivalent“ positions ...)

We use the metadata information provided by EBA

Template

F 01.01		Carrying amount
Balance Sheet Statement [Statement of Financial Position]: Assets		010
010	Cash, cash balances at central banks and other demand deposits	10898
020	Cash on hand	112718
030	Cash balances at central banks	151111
040	Other demand deposits	151115
050	Financial assets held for trading	11110
060	Derivatives	11108
070	Equity instruments	11112
080	Debt securities	11106
090	Loans and advances	11114
091	Trading financial assets	45921
092	Derivatives	45920

Use its metadata

Balance Sheet Statement [Statement of Financial Position]: Assets

		Carrying amount
		010
010	Cash, cash balances at central banks and other demand deposits	APL2576 ATY1177 BAS1506 MCY1881
020	Cash on hand	APL2576 ATY1177 BAS1506 MCY1878
030	Cash balances at central banks	APL2576 ATY1177 BAS1506 CPS1631 MCB5142 MCY2205
		APL2576 ATY1177 BAS1506

Code	Dictionary Label
APL2576	Accounting Portfolio „Cash“
ATY1177	Metric „Carrying Amount
BAS1506	Base „Assets“
MCY1881	Category „Cash on Hand“

e.g. for a the cell

010	
	APL2576 ATY1177 BAS1506 MCY1881

Show me the Metric „Carrying Amount“ (ATY1177) of all Accounting Portfolio „Cash on short notice“ (APL 2576) of Basetype „Assets“ (BAS1506) which are in the Main Category „Cash“ (MCY1881)

We use the metadata information provided by EBA

Template

F 01.01

Balance Sheet Statement [Statement of Financial Position]: Assets

		Carrying amount
		010
010	Cash, cash balances at central banks and other demand deposits	10898
020	Cash on hand	112718
030	Cash balances at central banks	151111
040	Other demand deposits	151115
050	Financial assets held for trading	11110
060	Derivatives	11108
070	Equity instruments	11112
080	Debt securities	11106
090	Loans and advances	11114
091	Trading financial assets	45921
092	Derivatives	45920

Use its metadata

Balance Sheet Statement [Statement of Financial Position]: Assets

Carrying amount

010

**Cells in FinRep~
10.000**

**Approx 300
members**

e.g. for a the cell

		010
		APL2576
		ATY1177
		BAS1506
		MCY1881

Show me the Metric „Carrying Amount“ (ATY1177) of all Accounting Portfolio „Cash on short notice“ (APL 2576) of Basetype „Assets“ (BAS1506) which are in the Main Category „Cash“ (MCY1506)

Mapping the Members

Cell-Information

	010
	APL2576
	ATY1177
	BAS1506
	MCY1881

Show me the Carrying Amount (ATY1177) of all Assets (BAS1506) which are of Type Cash (MCY1506) in the Accounting Portfolio „Cash on short notice“ (APL 2576)

Map each member

We: How can I find the Assets (Bas1506) in System A?

System: Please provide a Rule for finding all „Assets“

We: Here ist the rule!
Assets = „XYZ“ or „something else“. Please give us the keys which follow this rule

And get the keys

System A: Here are they keys

AccountNumbers:
{12345,
34567}

System B: Here are they keys

TradeIDs:
{A1,
A2,
A3,
A4}

Tagging the Data

Create a Big Table

Posting-ID	Hundred Other Dimensions	Carrying Amount
1	12345 A1	1
2	23456 A2	-2
...
1.394.023.344	34567 A3	100

Tagging according to the Rules

BAS	MCY	APL	GroupString
Assets	Cash	Cash on hand...	APL2576BAS1506MCY1881
Liabilities	something	else	APL2576BAS1508MCY1881
...
Assets	Loans	Amortized Cost	OTH9999KEY8888

Show me all **Assets** (BAS1506)

The cell as a search criteria

	010
	APL2576
	ATY1177
	BAS1506
	MCY1881

CatKey as Searchkey:

APL2576BAS1506MCY1881

Producing the Numbers

The final templates

Date		Report Date	Mandant								
Booking Date											
Currency Translation Date											
Cumulation Type	Not Cumulated										
Financial Report X	EF18.00										
Financial Report Y	EF18.00										
Mandant											
Posting Status	SENT										
Account Standard	IFRS										
To be compensated IFRS	(Members of to be compensated IFRS) - (Y)										
Ordinate Code	Ordinate Label	010		015		020		025		030	
		Amount / FX	Already Offset / FX	Amount / FX	Already Offset / FX	Amount / FX	Already Offset / FX	Amount / FX	Already Offset / FX	National / FX	Annex V.Part 2.70-71
	"Financial assets held for trading and trading"										
	"already offset"										
	"Financial liabilities held for trading and trading"										
	"already offset"										
010	"Interest rate"	-1.149.835.241,21	-1.862.063.301,75	929.591.138,94	2.158.546.959,72					143.833.577.758,15	
020	"of which: economic hedges"										
030	"OTC options"	-58.824.012,57	0,00	63.724.834,75	0,00					5.449.681.443,87	
030a	"OTC - credit institutions"	-58.604.195,06	0,00	62.324.814,78	0,00					4.957.446.340,06	
030b	"OTC - other financial corporations"			333.125,20	0,00					40.000.000,00	
030c	"OTC - rest"	-219.817,51	0,00	1.067.094,77	0,00					452.235.103,82	
040	"OTC other"	-1.091.011.228,63	-1.862.063.301,75	865.866.304,20	2.158.546.959,72					138.176.694.335,40	
040a	"OTC - credit institutions"	-980.518.319,48	0,00	769.508.192,10	0,00					24.834.121.187,46	
040b	"OTC - other financial corporations"	-4.424.826,42	-1.862.063.301,75	14.533.496,54	2.158.546.959,72					109.340.120.702,50	
040c	"OTC - rest"	-106.068.082,74	0,00	81.824.615,56	0,00					4.002.432.445,41	
050	"Organized market other"									207.201.379,87	
070	"Equity"	-5.482.214,57	0,00	5.508.958,16	0,00					446.384.628,53	
080	"of which: economic hedges"										
090	"OTC options"	-1.590.449,05	0,00	53.258,27	0,00					32.420.200,00	
090a	"OTC - credit institutions"	-1.590.449,05	0,00	53.258,27	0,00					32.420.200,00	
090b	"OTC - other financial corporations"										
100	"OTC other"			520.852,19	0,00					3.000.000,00	
100a	"OTC - credit institutions"			520.852,19	0,00					3.000.000,00	
110	"Organized market options"	-3.891.765,52	0,00	4.934.847,20	0,00					122.166.228,07	
120	"Organized market other"									288.798.260,48	
130	"Foreign exchange and gold"	-695.806.332,31	0,00	772.402.231,52	0,00					84.395.641.332,64	

Searchcriteria

SearchkeyA

Show me all „Derivatives“ of Risk-Type „Interest Rate Risk“ for „Financial Assets“ in the Accounting Portfolio „Held for Trading“.

SearchkeyB = A+ another Key

Show me all „Derivatives“ of Risk-Type „Interest Rate Risk“ for „Financial Assets“ in the Accounting Portfolio „Held for Trading“.

but only Instrument Type „OTC Options“

Testdata

Sidestep – some Experiments with GPT 4.0

GPT 4.0

C Now show instead of the categorisation keys the explanation according to the dictionary. Always start with the metric first.

AI Here's the table with explanations according to the dictionary. I've used Rust format

```

rust
+-----+-----+
|           | | "Carrying amount for Financial assets"
|           | | "Held for trading" ('APL2566ATY1177BAS1506')
+-----+-----+
| "Interest rate" | | In this cell there is the 'carrying amount' ('ATY1177')
| ('MCY1994TRI2701') | | for all 'Assets' ('BAS1506')
|           | | with the 'Accounting Portfolio' of type "financial assets"
|           | | but only namely for "Derivatives" ('MCY1994')
|           | | which are of "Type of Risk" "interest rate risk"
+-----+-----+
    
```

Erneut generieren

e.g. for a the cell

<h2 style="margin: 0;">F 10.00</h2> <p style="margin: 0;">Derivatives: Trading</p>		Assets
		010
010	Interest rate	APL2566 ATY1177 BAS1506 MCY1994 TRI2701

Show me the Metric „Carrying Amount“ (ATY1177) of all Accounting Portfolio „financial assets held for trading“ (APL 2566) of Basetype „Assets“ (BAS1506) which are in the Main Category „Derivative“ (MCY1994) of Risktype „Interest Rate“ (TRI2701)

Sidestep – some Experiments with GPT 4.0

GPT 4.0



Yes, it does help. Based on the Categorisation-Key

APL2566ATY1177BAS1506MCY1994TRI2701, the reported information for F10.00, r010, c010 is as follows:

- APL2566: Accounting Portfolio of type "financial assets held for trading"
- ATY1177: Carrying amount
- BAS1506: Assets
- MCY1994: Derivatives
- TRI2701: Type of Risk "interest rate risk"

So, for F10.00, r010, c010, the reported information is the carrying amount of financial assets held for trading, specifically for derivatives with interest rate risk.

e.g. for a the cell

F 10.00

Derivatives: Trading

Assets

010

010 Interest rate

APL2566
ATY1177
 BAS1506
 MCY1994
 TRI2701

Show me the Metric „Carrying Amount“ (ATY1177) of all Accounting Portfolio „financial assets held for trading“ (APL 2566) of Basetype „Assets“ (BAS1506) which are in the Main Category „Derivative“ (MCY1994) of Risktype „Interest Rate“ (TRI2701)

Sidestep – some Experiments with GPT 4.0

GPT 4.0

C what is the difference

The difference between F10.00, r010, c010 and F10.00, r010, c020 lies in the type of financial instruments they represent:

F10.00, r010, c010:

- APL2566: Accounting Portfolio of type "financial assets held for trading"
- BAS1506: Assets

F10.00, r010, c020:

- APL2566: Accounting Portfolio of type "financial liabilities held for trading"
- BAS1513: Liabilities

Both cells report the carrying amount (ATY1177) of derivatives (MCY1994) with interest rate risk (TRI2701), but one represents assets held for trading (c010) while the other represents liabilities held for trading (c020).

Comparing Cells

F 10.00 Derivatives: Trading		Assets	Liabilities
		010	020
010 Interest rate		APL2566	APL2566
		ATY1177	ATY1177
		BAS1506	BAS1513
		MCY1994	MCY1994
		TRI2701	TRI2701

Guidelines

Guideline 1

Before you think about correctness think about built in consistency

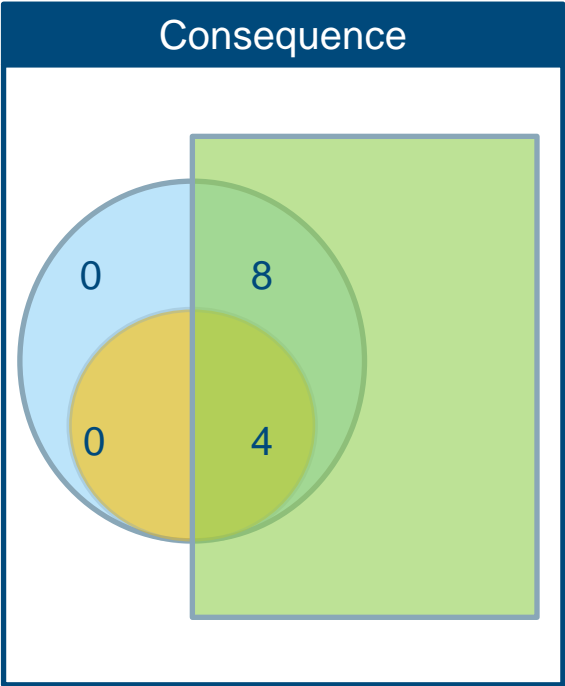
Use the build in datastructures within the taxonomies

DPM has built in Consistency Checks

Show me all „Derivatives“

... for assets

... but only Instrument Type „OTC Options“



Hierarchies, CatKeys, Dimensions and its members, appearances within the templates, cell definition (as intersection of a row and column).

relationship between members across dimensions, this „arghh-Dimension“ „MCY“ validations work only for numbers not for datasets and not just numbers

Validation for dataset

Guideline 1

Why validation for numbers
Aren't enough?

DPM has built in Consistency Checks

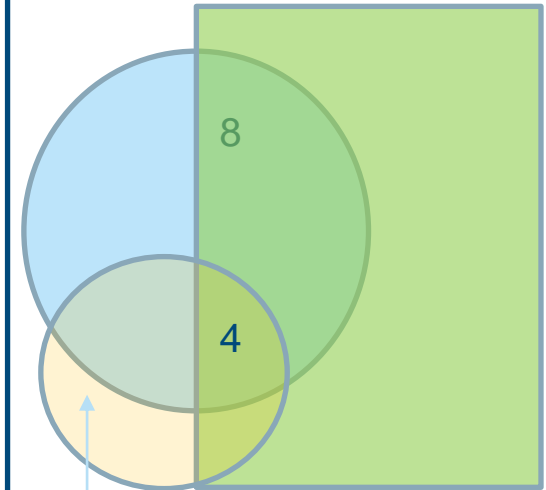
... Show me the assets

For all „Derivatives“

is a

... but only Instrument Type „OTC Options“

Consequence



This is a problem though numbers are the same and validation rules do still apply

Validation for dataset

Guideline 1

Why validating numbers
Is not enough?

DPM has built in Consistency Checks

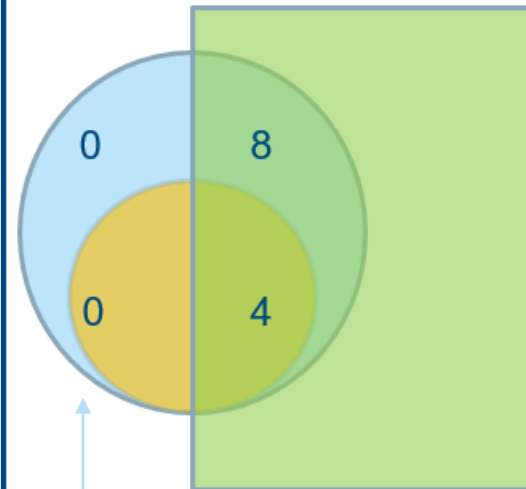
... Show me the assets

For all „Derivatives“

is a

... but only Instrument Type „OTC Options“

Consequence



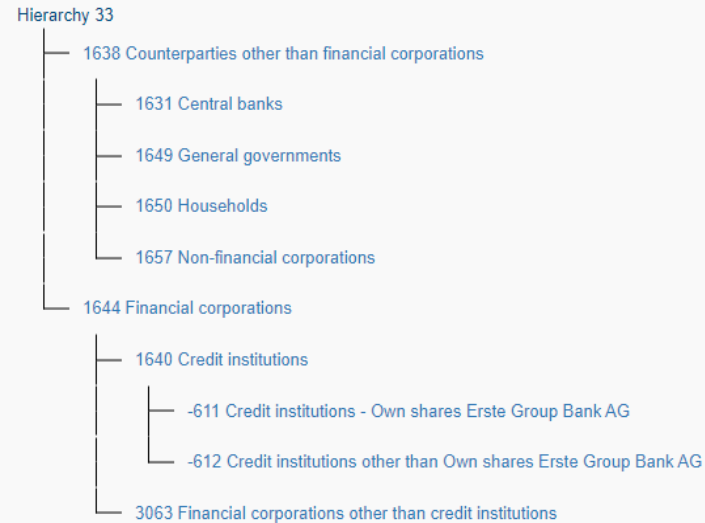
It should look like this because OTC Options is a subset of „derivatives“

Guidelines

Guideline 2

Add words to the dictionary if the meaning is not precise enough

Example



Guidelines

Guideline 3

Go for the meaning and find every **meaningful** error where **consistency is guaranteed**:

Meaningful means errors above 2€

Example

Definition: $C = A + B$

Option 1)
 $C = A \cup B$

Option 2)
A is mapped to data
B is mapped to data
C is mapped to data

⇒ Weird technical problem

⇒ Additional analyses

Advantage 1: Metadata gives Data a meaning

DPM as a Metadatamodel

... DPM Metadata provides meaning to data e.g. we know exactly what the number is composed of.

Other Metadatamodels

DPM provides this, without tagging or doing anything AFTER the data is prepared. It could be done before!

Sidestep – some Experiments with GPT 4.0

GPT 4.0



can you guess what the connection between MCY46166 and MCY46167, MCY46168 and MCY 46169 is

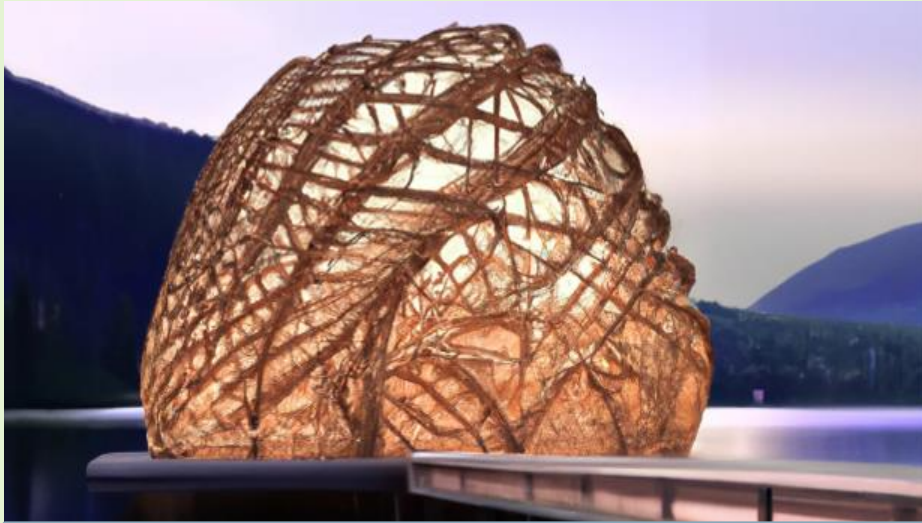


Based on the member labels, it seems that MCY46166 represents the total amount of cash and cash balances at central banks and other demand deposits, while MCY46167, MCY46168, and MCY46169 represent the individual components that make up this total amount. MCY46167 represents the amount of cash on hand, MCY46168 represents the amount of cash balances at central banks, and MCY46169 represents the amount of other demand deposits.

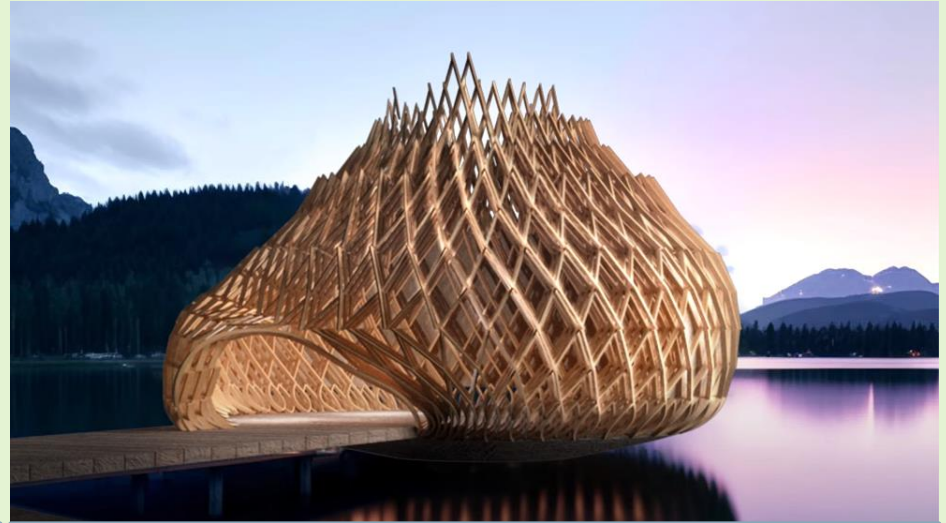
Wrong answer

- Cash and cash equivalents
 - Cash on hand
 - MCY46167
- Cash balances at central banks
 - MCY46168
- Other demand deposits
 - MCY46169
- Cash and cash balances at central banks and other demand deposits
 - MCY46166
 - APL2576 -> 2576

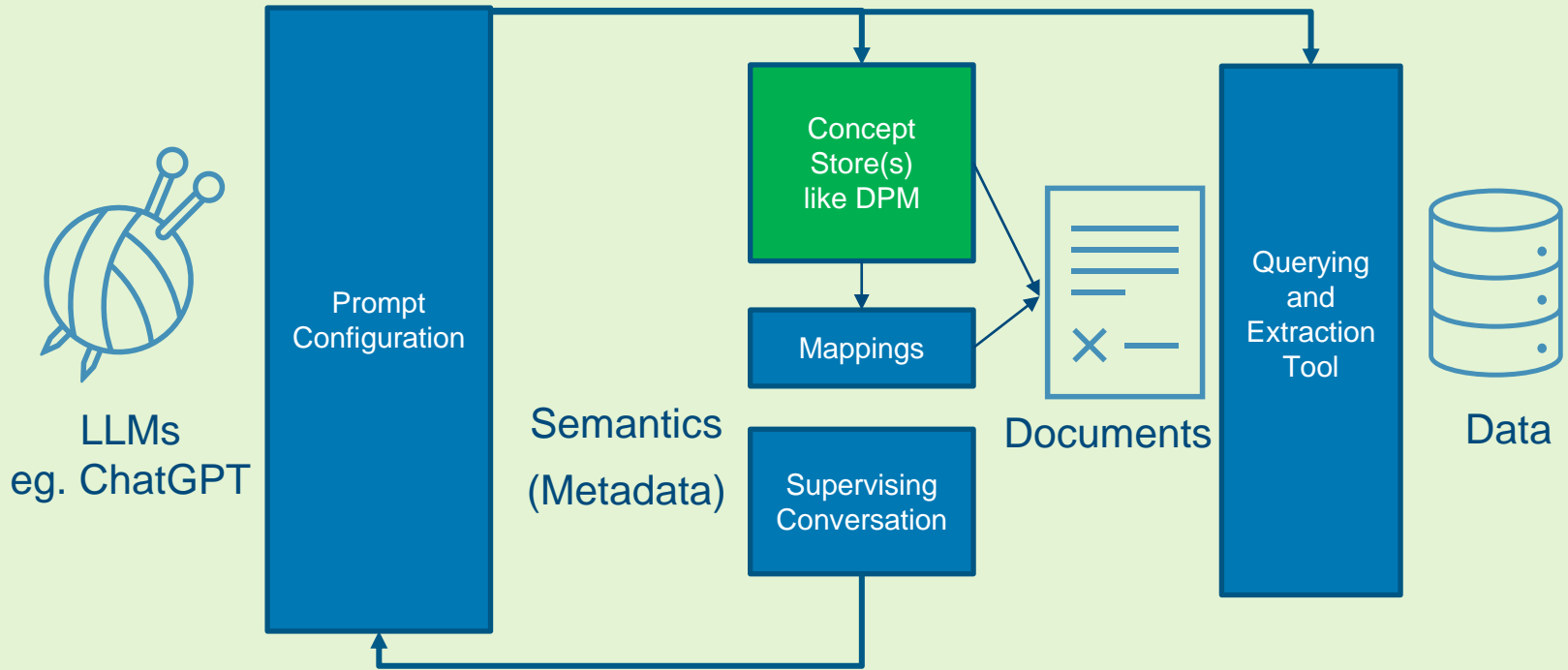
How we want to use LLMs and other tech



„Conversational Intention
Recognition Machine“
like LLM



„Computational
Knowledge Engine“
Like Wolfram



Disclosures in accordance with EU Taxonomy: The key performance indicators in this section were determined based on Commission Delegated Regulation (EU) 2021/2178 in conjunction with the International Financial Reporting Standards applicable for the Consolidated Financial Statements. Thereby, revenue, capital expenditures and operating expenses were predominantly directly allocated to an economic activity listed in Delegated Regulation (EU) 2020/852; in determining capital expenditures and operating expenses, allocations were also made based on the revenue of the Taxonomy-eligible activities. To avoid double counting, the allocation was always made to one economic activity only. **Taxonomy-eligible revenue accounted for 20% of revenue** according to the Consolidated Statement of Income in the reporting year. In the reporting year, Taxonomy-eligible capital expenditures accounted for 40% of additions (including additions from business combinations) to other intangible assets and property, plant and equipment in accordance with Note 13 to the Consolidated Financial Statements. Taxonomy-eligible operating expenses accounted for 14% of the corresponding expenses recognized in the Consolidated Financial Statements in the reporting year. The remaining portions of the key performance indicators are not Taxonomy-eligible. Our main Taxonomy-eligible economic activities are derived from the manufacture of low-carbon transport and energy-efficient building technologies (mainly from Mobility and Smart Infrastructure operations), transport infrastructure (from Mobility operations) and the service of energy-efficient building technologies (from Smart Infrastructure operations), as well as the Group's own real estate portfolio. The majority of Taxonomy-eligible capital expenditures result from the latter economic activity. The above-mentioned economic activities refer to chapters 3, 6 and 7 of Annex I of Delegated Regulation (EU) 2020/852.

Some reports text only

Real world examples

Turnover		2021/22
Turnover (= denominator of KPI)	EURm	4,062.2
thereof taxonomy-aligned (= numerator of KPI)	EURm	1,105.8
Turnover KPI	%	27.2

CapEx		2021/22
Additions to intangible assets, fixed assets and rights of use (= denominator of KPI)	EURm	572.1
thereof taxonomy-aligned (= numerator of KPI)	EURm	484.4
CapEx KPI	%	84.7

Many create their own layout

And other issues
(language, cut offs,...)

**We want to
share this data**

Explored Benefits

Errors are easier to find

Changes are much easier to introduce

Easy Code maintenance (metamodel contains most of the information)

Validation follows meaning

Insane increase in Dataquality (by following the guidelines)

OOB maintenance by the regulator

DPM is agnostic to models and formats

Analytics included

Separation Concern (Semantic Rules, Data)

Yet Unexplored Potentials

universal „conversational
data retrieval interface“
(versus linear Chatbots)

Creating a
marketplace for
data

Don't Search – Ask!
which creates improvable
knowledge

Connecting different
models via
metadatamodels

Sidestep – some Experiments with GPT 4.0 regarding connecting different data models

OeNB Wiki

FUNKTION XEN_Sektor_FinRep(AI_Mandant MA, AI_Einheitennummer_ID emld, AI_Stichtag_Datum repDate)

esvg_sektor = SELEKTIERE EM.EMA49_Sektor_fuer_Meldezwecke_Code VON EM_Einheit_MS

internat_org = SELEKTIERE EO.EO40_Internationale_Organisation_OS_Code VON EO_Einheit_OS VON EM_Einheit_MS

kl_gem_finrep = SELEKTIERE EM.EMA41_Kreditinstitut_gem_FinRep_Kennzeichen VON EM_Einheit_MS

WENN(esvg_sektor ISTGLEICH "Zentralbank (1210)" UND NICHT(internat_org ENTHAELT ("Bank für Internationalen Zahlungsausgleich (5B)", "Internationaler Währungsfonds (1C)")) DANN

rValue = "Central banks (1)"

SONST WENN(esvg_sektor ENTHAELT ("Zentralstaat (1311)", "Länder (inkl. Landeskammern, Landesfonds) (1312)", "Gemeinden (inkl. Gemeindefonds und -verbände) (1313)", "Sozialversicherung (1314)", "Staat nicht zuordenbar (1300Z)") ODER internat_org ENTHAELT ("sonstigeInternationaleOrg_GR")) DANN

rValue = "General governments (2)"

SONST WENN(kl_gem_finrep ISTGLEICH WAHR) DANN

rValue = "Credit institutions (3)"

SONST WENN(esvg_sektor ENTHAELT ("sonstigeFinanzUnternehmen_GR") UND internat_org ISTGLEICH NULL) DANN

rValue = "Financial corporations other than credit institutions (4)"

SONST WENN(esvg_sektor ISTGLEICH "Nicht-finanzielle Unternehmen (1100)") DANN

rValue = "Non-financial corporations (5)"

SONST WENN(esvg_sektor ENTHAELT ("Selbständigenhaushalte (mit u. ohne Arbeitnehmer) (1400A)", "sonstige private Haushalte (1400B)", "private Haushalte nicht zuordenbar (1400Z)", "private Organisationen ohne Erwerbszweck (1500)") DANN

rValue = "Households (6)"

ENDE

RUECKGABE rValue

DPM Description

F 10.00 

Derivatives: Trading



Assets

Liabilities

010

020

010

Interest rate  

APL2566

ATY1177

BAS1506

MCY1994

TRI2701

APL2566

ATY1177

BAS1513

MCY1994

TRI2701

Sidestep – some Experiments with GPT 4.0 regarding connecting models

GPT 4.0 Conversation

C which dictionary entries do you know and can you please provide the descriptions

Here is a list of dictionary entries (Categorisation-Keys) and their descriptions that I have encountered in our previous discussions:

1. ATY1177: Carrying amount (Netto-Buchwert)
 - * This represents the carrying amount of financial instruments.
2. MCY1994: Derivatives (Derivate)
 - * Financial instruments whose value is derived from the value of an underlying asset, index, or reference rate.
3. TRI2701: Interest rate risk (Zinsrisiko)
 - * The risk that changes in interest rates will adversely affect the value of a financial instrument.
4. TYH512: Economic hedges (Wirtschaftliche Sicherungen)
 - * Hedging instruments that are used to manage the exposure to interest rate risk.
5. TMA1822: OTC options (OTC-Optionen)
 - * Options that are traded over-the-counter, rather than on a regulated exchange.
6. MCB5131: OTC interest rate options (OTC-Zinsoptionen)
 - * Interest rate options that are traded over-the-counter.
7. APL2566: Financial assets/liabilities held for trading (Finanzinstrumente zum Handel gehalten)
 - * Financial assets or liabilities that are held for the purpose of short-term trading, with the intention of generating profits from fluctuations in market prices.
8. BAS1506: Assets (Aktiva)
 - * Resources controlled by an entity that are expected to generate future economic benefits.
9. BAS1513: Liabilities (Passiva)
 - * Obligations of an entity that it needs to settle in the future.

Please note that this list is not exhaustive, and there may be other relevant dictionary entries in the context of financial reporting. However, these are the main ones we have discussed so far.

DPM Description

F 10.00 Derivatives: Trading		Assets		Liabilities	
		010	020	010	020
010 Interest rate		APL2566	APL2566	ATY1177	ATY1177
		BAS1506	BAS1506	BAS1513	BAS1513
		MCY1994	MCY1994	MCY1994	MCY1994
		TRI2701	TRI2701	TRI2701	TRI2701

Yet Unexplored Potentials

(1)
Creating a universal
„conversational data
retrieval interface
(versus Chatbots)

(3)
Creating a
marketplace for data

(5)
Centralizing meaning
and using distributed
data

(7)
Connecting data
(metadata, data,
taxonomies, reports)

(2)
Don't Search – Ask!
which creates
improvable
knowledge

(4)
Connecting different
models via
metadatamodels

(6)
Self Tutoring and
Learning

Potential 7: connect FinRep and IFRS Taxonomy and self tagging!

Challenge

Reports should be machine readable and comparable. While machine readability is not a problem, comparability is a hard task?

LLM and Metadatamodel

...comparability gets possible even if different words or tags are used? What if words and data is connected, not via html links but via an ontology.

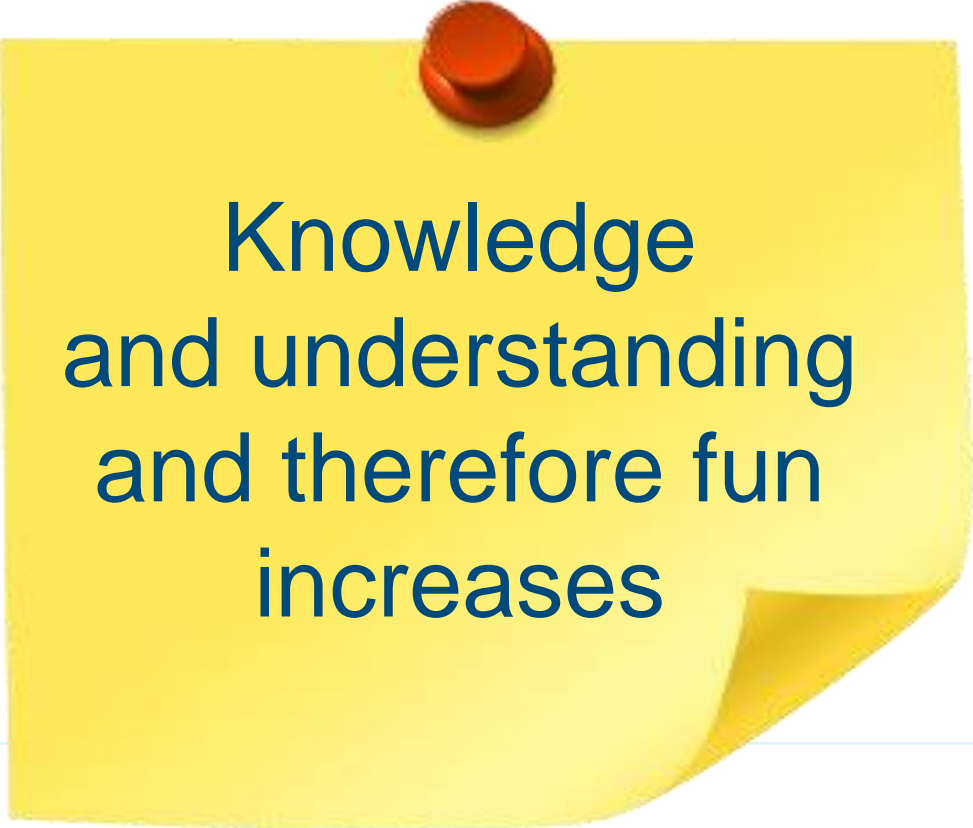
Technical Feasibility

IFRS Taxonomy

ESEF

FinRep Taxonomy

Explored but completely underestimated benefit



**Knowledge
and understanding
and therefore fun
increases**

**“... LET THE MACHINES GET
ON WITH THE ADDING UP,”**

**“(No)...WE DEMAND RIGIDLY
DEFINED AREAS OF DOUBT
AND UNCERTAINTY!”**