

PROJECT: FUNDS OFFICIAL DATA REPOSITORY

SUMMING UP

OBJECTIVE

We need a single repository for OMAM funds official return data. Several fund characteristics can be part of the data model design. This will allow us to tailor our answers depending on the angle the questions will be asked.

This project will provide various stakeholders with unique answers to questions related to official funds performances and characteristics.

BENEFITS

Plenty of MS Excel spreadsheets are currently used to produce our external and internal reporting. They constitute a big source of operational risk and they won't be needed anymore.

Tasks, responsibilities and validation processes will be correctly identified

Related procedures will be centralized, correctly structured and maintained.

CHALLENGES

The complexity of the project is linked to:

- Collecting data from the past (ideally we need to go back to 2004).
- Making sure all stakeholders have their needs covered.
- Identify roles and responsibilities for the initial load, update, maintenance and validation of each specific data subset.
- Obtain from all the stakeholders the relevant specifications of the rules we need to use the data we collect (e.g. from net returns to gross returns, HWM calculations).
- Correctly deal with the logic linked to the production of the returns time series (e.g. dealing with funds merge/spin off¹).
- Provide efficient download/upload procedures from the selected data providers/to relevant data vendors/clients platforms
- Manage dates/versioning: roles change, objectives change, peer group and benchmark references can change. We have to accommodate for this.

¹ As an example with UKDEFOS replacing UKDEF we should be able to link past UKDEF performance with UKDEFOS one otherwise we will have to wait a long time before being able to provide a decent time series of returns.

SCOPE

The scope of the project can grow according to the characteristics we want to model and capture.

Although the main objective for this project is to be able to provide funds returns, NaVs and performance against objectives (benchmarks, peer groups), this objective can be expanded including:

- References to Fund Managers roles and significant dates
- Client custom classifications

We are interested in the official data only and not in how they are produced. This is definitely not about portfolios constituents/positions and valuations.

When the fund unit price has been collected and stored, we will assume the basic validation about positions and trades has been already performed. The only thing we are interested into is to make sure that this number is consistent with what fund performance estimates (as provided by Factset or Vivaldi as an example) were telling us.

RISKS

There are several risks to this project because of:

- Potentially (very) wide scope
- Several stakeholders involved
- Various data sources
- Potential confusion about data validation issues
- Added burden on some OMAM functions
- Clashes with other projects and potentially creating duplicates inside OMAM hub. While avoiding this is a commendable task I believe being too fanatic about it might lead to big delays: entire lines of work (e.g. Vivaldi) might need to be refactored if we want to be perfect from this point of view. Nevertheless this is something we need to keep in our minds.

DETAILS

OBJECTIVES

OMAM products are the funds we sell.

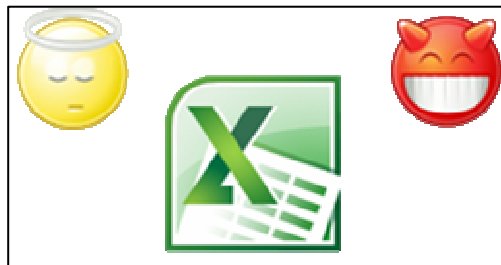
Our clients, on top of good performances, expect from us quick and precise answers to all their questions concerning the performances of their investments.

OMAM management needs timely, precise updates about our fund performances and their economic impact on the company results.

Fund performances and all the ancillary financial calculations can be distributed using several means (XL spreadsheets, PDF reports, PPT presentations, emails) but the production of the underlying calculations needs to be provided by a centralized source.

Using spreadsheets and different data repositories is:

- 1) **Inefficient and potential source of errors:** because we end up replicating the same set of data multiple times and out of them we might end up with two different answers to the same question.
- 2) **Source of operational risk:** because we end up relying on individual efforts and not on established, documented, updated and tested procedures. MS Excel spreadsheets just can't be the basis of correct enterprise procedures: they are intimately linked to their user/creator and their sharing is always problematic. This is an obvious source of *key man* risk.
- 3) **Source of reputational risk:** beside the obvious TCF concerns, being wrong or late could really impact our business negatively.



The relevant story here is: we need a single data source when we try to represent our funds performances.

That has to be the basis of all our reporting and from there we have to upload to the other data vendors or external systems (e.g. Factset/SPAR) when this is relevant.

THE DATA MODEL

The technical basis of the project is to identify the right way of representing the data source underlying all the processes that will be run in order to achieve the various objectives.

Once this is done we can then split that in various “areas” and allocate tasks and responsibilities accordingly

This is the exercise that Praveen and I have run during the past month. Please refer to Praveen document for the most updated details.

What follows is presented only as a helper. I'd like to provide some visual about what we believe are the most relevant set of information we will need to store in order to provide the essential portion of the project.

These diagram are not necessarily in line with the latest version as presented by the *Requirements Specification Document*. That is the relevant IT document that will represent the blueprint of the dBase underlying the project.

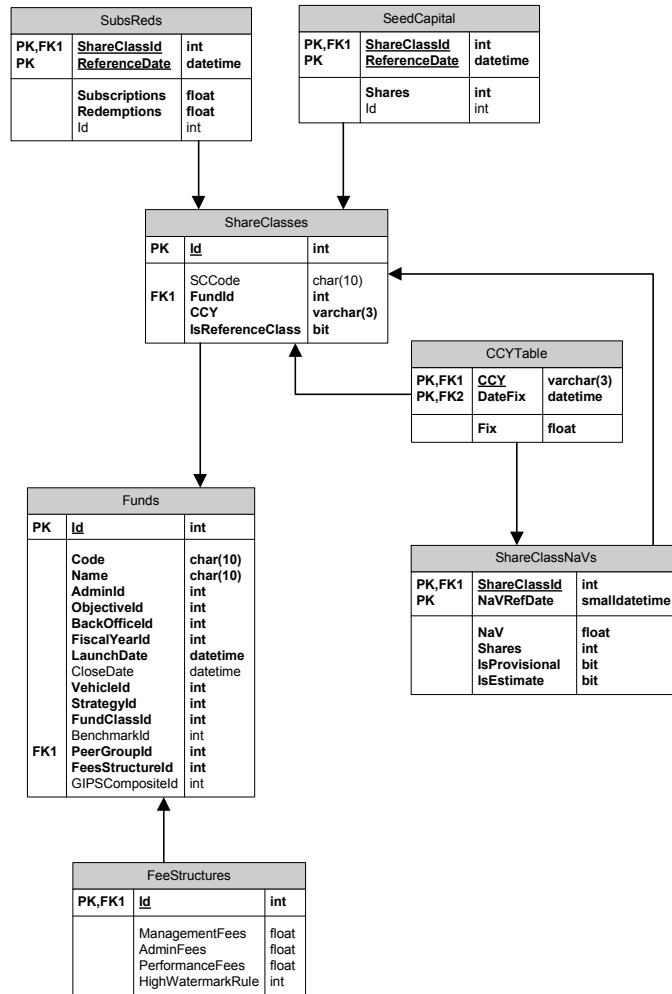


Figure 1: core dataset

The core of the database (Figure 1) is based on the notion that each fund Value can be represented as the sum of all the share classes' prices time number of shares².

² For segregated accounts the concept of share classes and number of shares is redundant. But this structure is still valid if we say number of shares=1 for a unique share class. Using a very generic notation we can say:

$$FundsValue = Sum(ShareClasses.Shares \times ShareClasses.NaV \times ShareClasses.FXconversion)$$

where $ShareClasses.Id \in Funds.Id$

The performance of a fund will be associated with the evolution of the NaV of a selected share class (*ShareClasses.IsReferenceClass = TRUE*)

Once we get this we will be able to retrieve and distribute the time series of returns for all our funds and we will be able to answer question about period returns (gross, net³), asset under management, fees impact, fund profitability.

Another relevant byproduct would be the automation of the GIPS composites construction.

Notice as well that, thanks to the *IsProvisional* and *IsEstimate* flags in the crucial ShareClassesNaVs table, we will be able to provide those items in a timely fashion based on provisional and/or estimated numbers.

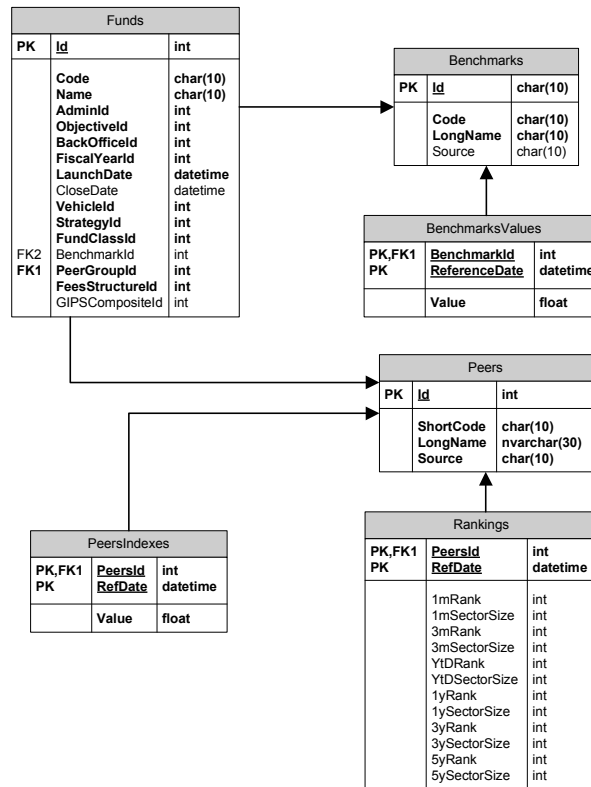


Figure 2: relative performance data

The second essential bit of the database (Figure 2) deals with benchmarks and peers performance data.

The first thing is to define the objective of each fund: this can be *beat a benchmark*⁴ or *beat a peer group*. There is nothing in the way of measuring each fund performance against both.

³ There is a whole bunch of fees that should be considered: (management, performance, administrative, custodian). On top of this there are plenty of details about caps, floors, hurdles that we might want to consider depending on the degree of precision we want to achieve. Technically there are expenses as well that might be modelled but I'm not sure we want to go down that route.

⁴ We have to be careful with those funds that need to beat (benchmark + spread). Libor benchmarks need to be created as well for absolute return funds and for all the hurdle references (crucial for gross/net calculations).

Getting this right means we will be able to automate the production of (as examples):

- Green stripy (weekly)
- MI performance reports
- Traffic lights
- TCF tables

This is on top of being able to answer all the questions about our range of product performance as an aggregate (e.g. percentage of AUM better than objectives).

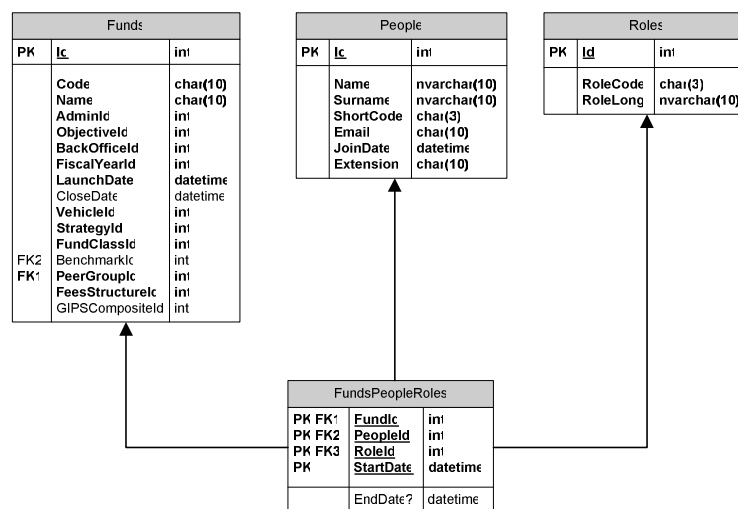


Figure 3: fund roles

This third block (Figure 3) merges fund details and performance with the people running them. If we get this set of information we will be able to answer question like:

- What was the performance of this fund since the fund manager XY took over?
- How much money is a specific desk/fund manager managing?
- What is the performance of funds belonging to a specific desk?

OUT OF SCOPE

One of the main dangers of the project is overextending its scope. It is very easy to try to fit in all you can think about and, honestly, there are a lot of things that potentially can be part of this.

One thing being mentioned, as an example, is the reference to specific client classifications. While this is a *nice to have* I do not think it should be part of the initial effort.

What is really crucial instead is to have a good idea about what could be the potential project ramifications and structure the data model accordingly.

VALIDATION

Validation plays a crucial role. We want a unique source of data but we need to make sure it's right: being always wrong is not better than being somewhere right and somewhere else not so (like it can happen right now).

This project is not about how periodic NaVs are produced: no validation of portfolio positions, market prices, accruals, corporate actions is involved here.

The right place to run all those reconciliation processes is the Ops desk. And this project has nothing to add there. Beside *thinkFolio* is the right application where all the fund managers can double check their portfolios.

The process starts once we receive the NaV numbers from the relevant providers (RBC, Citi, GlobeOp). The Ops team is our first line of “defence” and ideally the first check about those numbers will be performed there⁵.

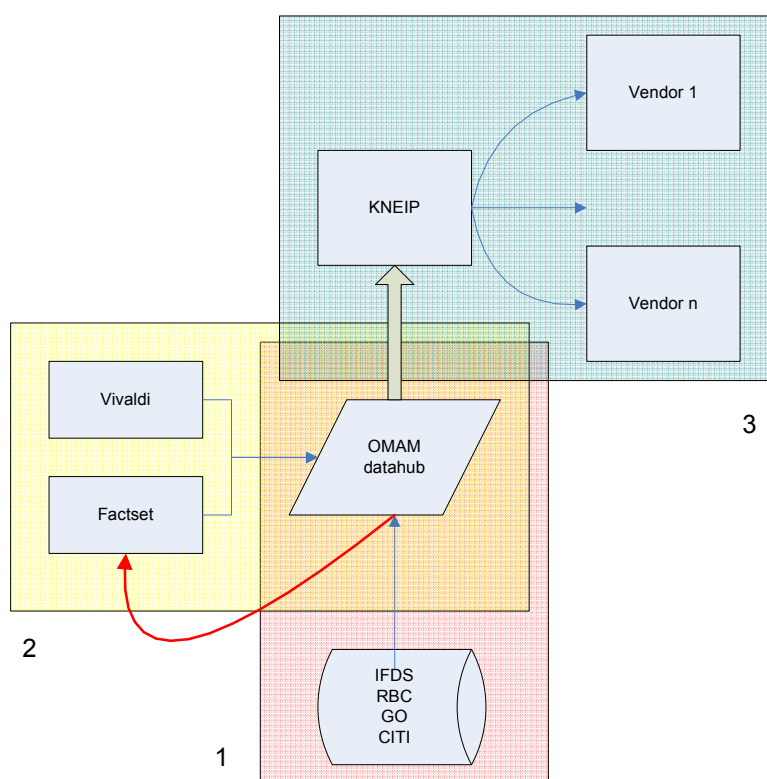


Figure 4: data validation

Once Ops team uploads them (item 1 on Figure 4) there is a potential second line of “defence” provided by OMAM sources of fund performance estimates: Factset and Vivaldi (item 2 on Figure 4⁶).

⁵ I’m thinking along the lines of a sense check. If the number is abnormal (we can set some sort of thresholds) Ops might raise questions.

⁶ The red line refers to the upload from the *datahub* to *Factset/SPAR*

Potentially there is an interesting part of validation exercise we might want to perform and this project might help somehow: the number of share as being provided by the TPA is currently not verified beyond making sure that the fund Value (i.e. shares x price) fluctuation is within a certain threshold.

Technically there will be the option to check that number based on the subscriptions/redemption records.

Thanks to Kneip we have outsourced all the checks with third parties data vendor that publish our NaV numbers and fund details (item 3 on Figure 4).

I believe there is very little value into adding a reverse arrow providing validation of our data from Kneip.

TASKS AND RESPONSIBILITIES

Splitting the database in different areas makes easier to identify tasks and responsibilities. This is a quick outline:

Operations

- 1) Gathering the NaV numbers as supplied by the different providers
- 2) Validate those data both against generic thresholds (i.e. there is a fair chance that a 10% performance on a fixed income fund might be wrong) and early estimates (Vivaldi estimates are available at around 5pm same day, *Factset* estimates are available early in the morning the following day).
- 3) Uploaded data to the dBase

Risk & Performance

- 1) Provide all the updates for peer group performances (this is done on a weekly basis)
- 2) Provide all the updates for benchmark performances (ideally we could automate that by using the Bloomberg data feed)
- 3) Specify and maintain the GIPS composites
- 4) Define what has to be uploaded to Factset/SPAR database
- 5) Provide *Factset* estimates of P&L to Ops for their validation process
- 6) Provide Vivaldi estimates of funds P&L to Ops for their validation process

Human resources

- 1) Keep track of the different roles played by OMAM employees as relevant for for each fund. This is basically making sure that what on Figure 3 is always updated.
- 2) Define what is relevant: Fund Managers, Head of Desk, Deputies seem an obvious choice. Should we include analysts as well?

I understand that going back to 2004 and recompile all the relevant information (even for dead funds!) is quite daunting task but this is needed in order to be able to provide answers to all the questions we are asked. People, Fund, Role and dates are needed.

Sales & Marketing

- 1) Make sure all the fund characteristics (benchmarks, peers group among others) are correctly specified and maintained going forward.
- 2) Specify each fund objectives (peer groups, benchmarks)
- 3) Timely update that whenever there is a change
- 4) Provide feedback from client and potential clients

Finance

- 1) Identify all the fee structures by fund and maintain the relevant dBase
- 2) Specify all the relevant rules needed to calculate fees, gross returns
- 3) Maintain and create the dataset about fund start/end dates and fiscal year rules

Sales & Marketing support

- 1) Specify factsheets requirements
- 2) Specify clients requirements by funds
- 3) Check the feed to Kneip

IT

- 1) Implementing the dBase design
- 2) Collecting all the relevant specifications for the different rules that have to be implemented
- 3) Implement all those rules
- 4) Provide all the relevant stakeholders with the appropriate interfaces to perform all the Create/Read/Update/Delete operations
- 5) Implement and deliver all the relevant reports as specified by their end users
- 6) Maintain all the logs about who did what and when
- 7) Centralize all the relevant procedures belonging to the relevant activities around the database.