



# 金融資安研習營

資料倉儲-資料治理

張皇裕 Henry



捷智商訊科技

人才、專業、責任、以客為尊



# Henry 資歷



- **Henry**
- **金融科技資歷：22年**
  - 捷智商訊 RD部門 副總
  - 凱基銀行 數金處
  - 開發金控 資訊處
  - 台新銀行 資訊處
- 主要經驗在資料倉儲、數位通路領域。
- 負責過台新銀、開發金資料倉儲建置，及行銷自動化系統建置專案。
- 負責過數位帳戶Richart系統建置專案



# 簡報大綱

**01.資料倉儲系統架構簡介**

**02.資料治理實務管理方式**

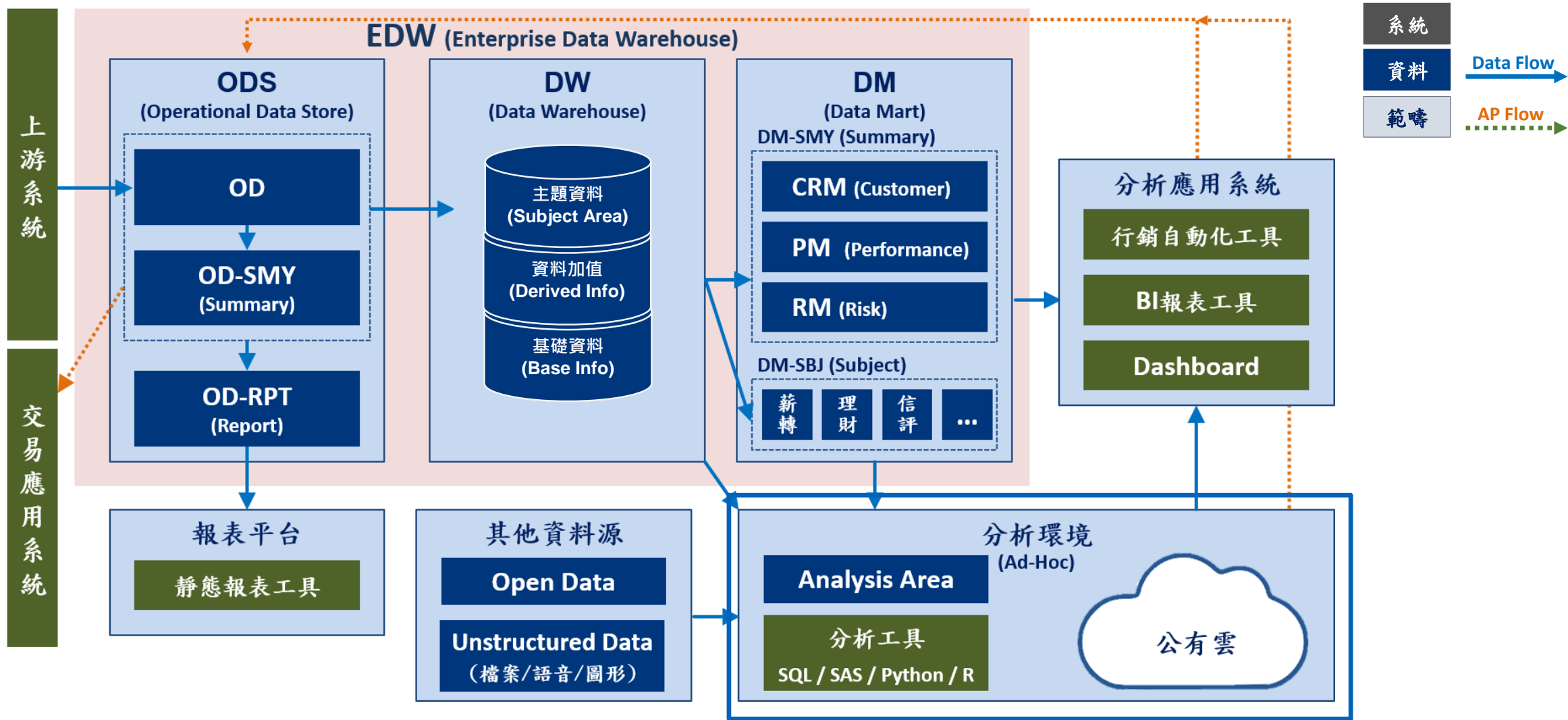
**03.總結**



# 資料倉儲系統架構 – 資料板塊



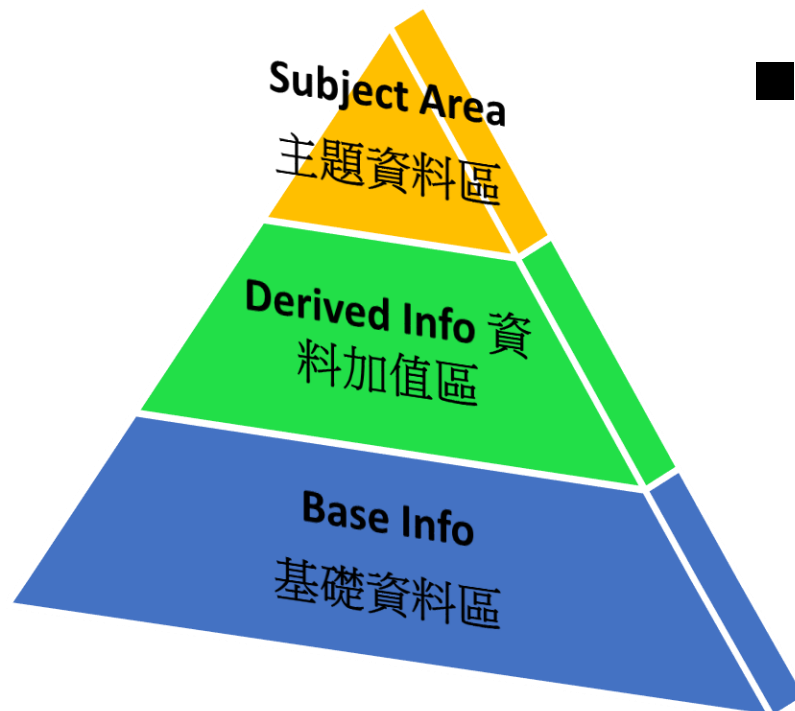
# 資料倉儲系統架構 – 資料流程與應用



# 資料倉儲系統架構-DW資料模型設計概念

## ■ Base Info (基礎資料區)

- 整合企業資料
- 資料範例
  - Customer
  - ✓ Account
  - ✓ Card
  - ✓ Transaction



## ■ Derived Info (資料增值區)

- 基礎資料經邏輯運算後產生新的資訊
- 資料範例
  - ✓ AUM
  - ✓ 平均餘額
  - ✓ 近一個月存提款次數

## ■ Subject Area (主題資料區)

- 根據特定議題彙整基礎及增值區資料
- Dimensional Modeling Techniques
- Data Model
  - ✓ Dimension Table
  - ✓ Fact Table

# 資料倉儲系統架構-DM資料模型運用

## ■多樣性的分析主題

### 業績分析

營業員產品績效

前N大營業員績效

營業員產品客戶通路績效

### 營運分析

營業單位產品績效

通路績效

目標達成率

營績效獲利虧損

利潤中心經營成本分析

市佔率波動分析

### 客戶屬性分析1

新客戶開發

客戶貢獻度分析

客戶RFM分析

前N大高風險客戶

客戶財務狀況分析

客戶分群profile分析

### 客戶屬性分析2

客戶交易分群分析

前N大帳戶交易分析

銷售通路分析

### 產品分析模型

各產品目標達成率評估分析

產品績效獲利虧損分析

各產品市佔率波動分析

各產品之客戶profile分析

### 財務分析

公司經營績效獲利虧損分析

利潤中心經營成本分析

### 組織分析

產品績效

通路績效分析

目標達成率分析

獲利虧損分析

經營成本分析

### 風險控管分析

績效獲利虧損

客戶交易行為

前N大帳戶交易

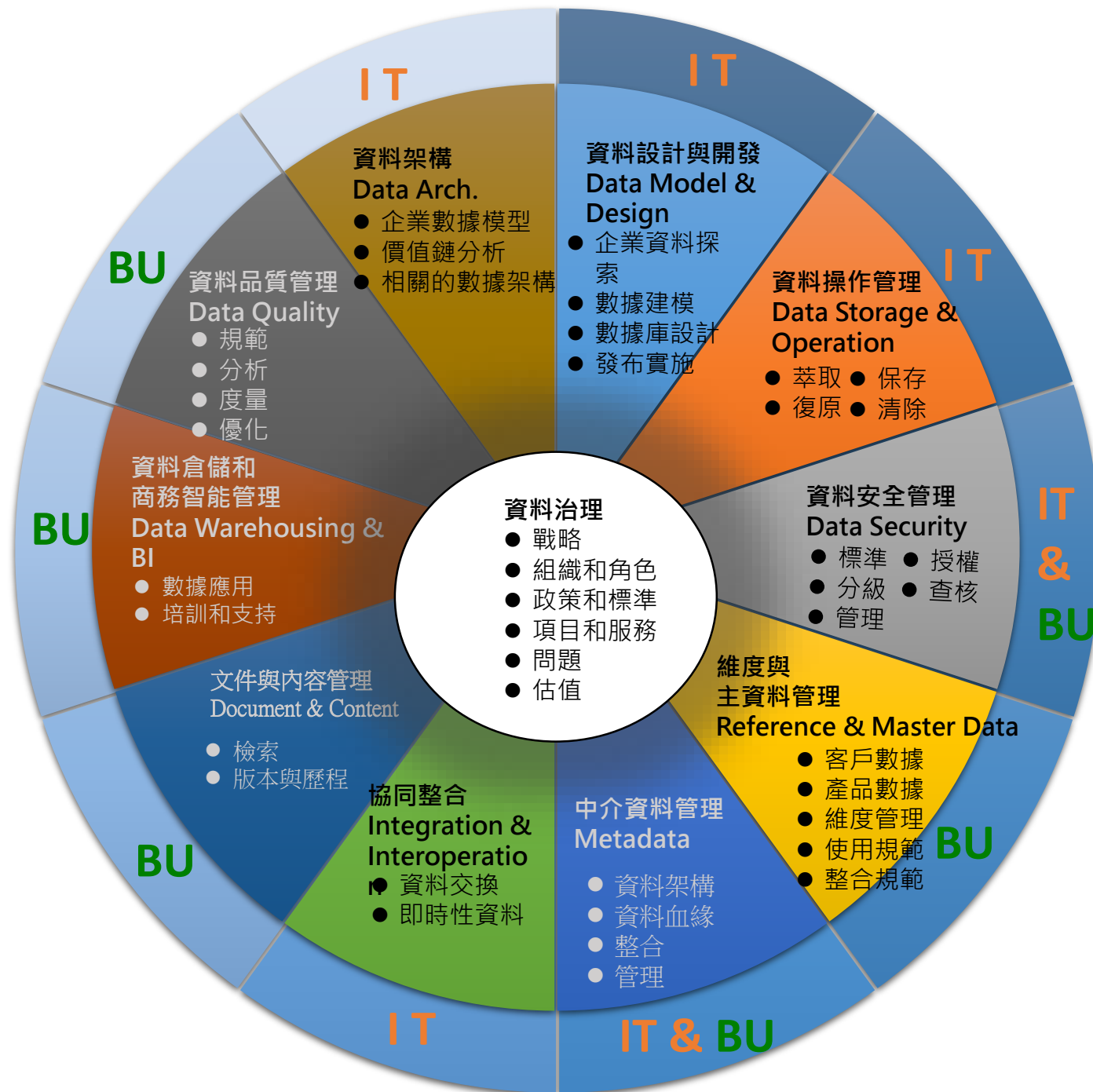
風險客戶

異常交易分析

交易額度分析

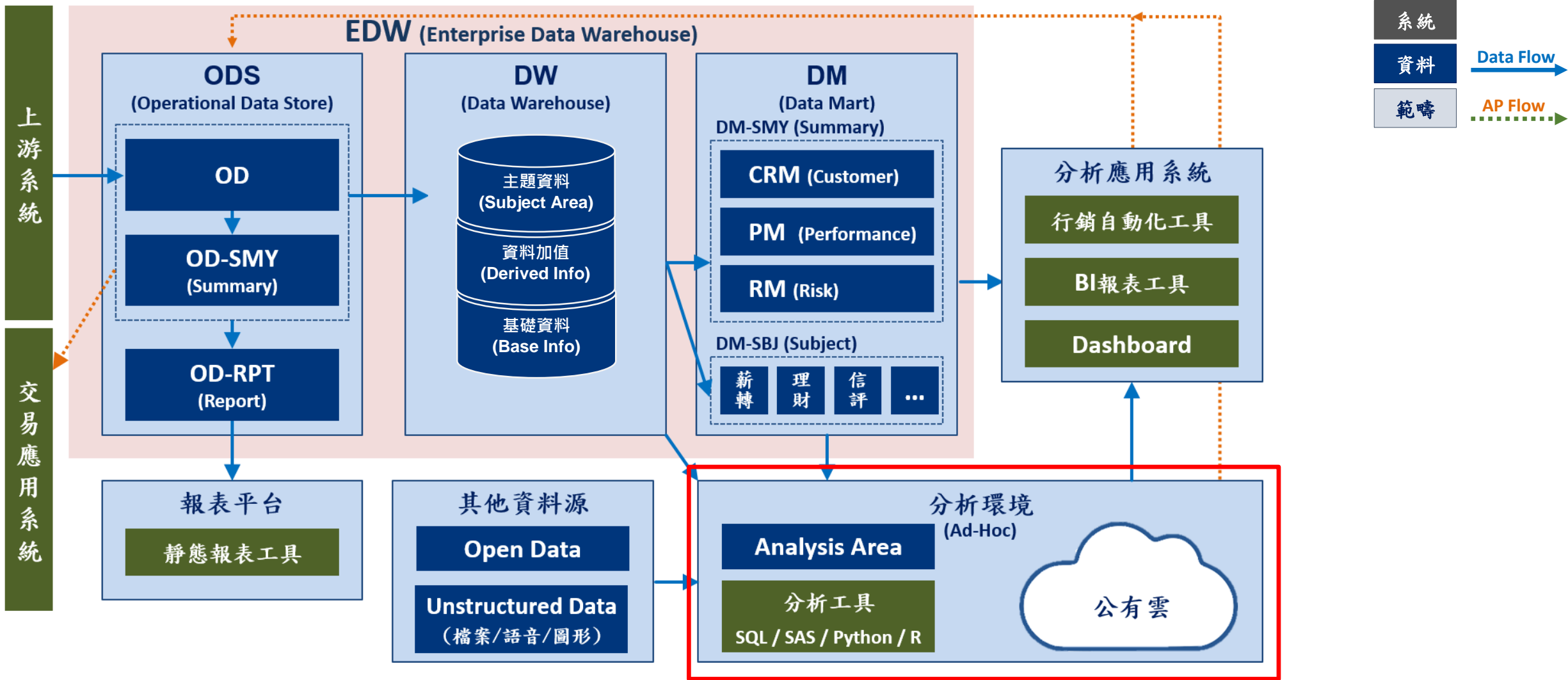
錯帳比率分析

# 資料治理藍圖





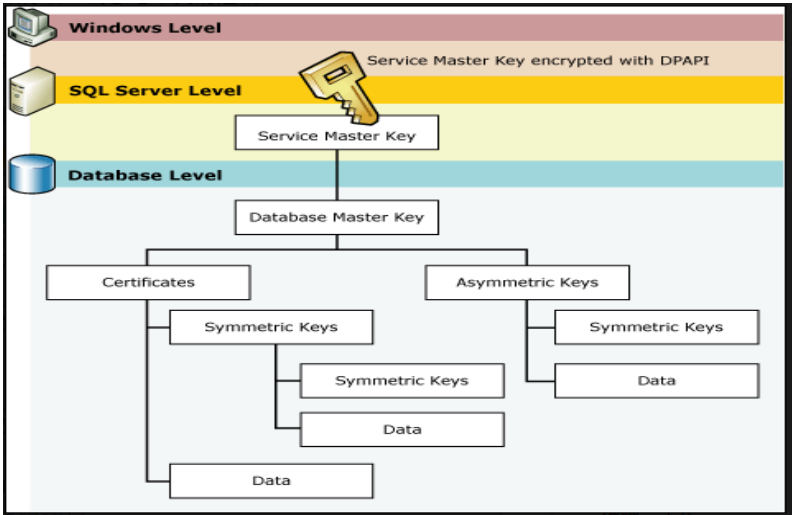
# 資料倉儲系統架構 – 個人資料保護與管控



# 資料治理 - 資料Key值加密

- 重要客戶資料同時也是資料Key值，ID、帳號、卡號
- 兼顧個資資料保護與資料分析自由度
- 常用Key去識別化方法：查表法、DES加密法....
- 可解密還原、長度短、純文字、唯一性

序號	ID
1	A167871668
2	A115785879
3	A109164935
4	A185807424
5	A171105220
6	A154901326
7	A105243757
8	A113216862



身分證/護照	帳號
A123456789	0021234567890

- 資料加密
  - 金鑰(對稱/非對稱)
  - 憑證
  - 函數

身分證/護照
0x010000003B42DEA950744CA93935C5A05C49F18CAF070FE247CE1A10C27A7E731657085F02FE19DA26830A60
帳號
0x010000002BB9511035F2DF6B9F7C7F4AB196BAD2F1F28523B869B38BAB0EF5ABB754F44666FBDDF2F4D64801

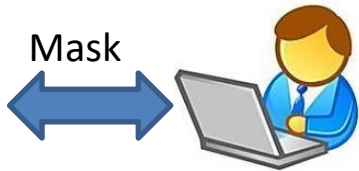
# 資料治理 – 個人資料去識別化

- 客戶資料遮罩，兼顧資料保護與資料分析價值
- 依資料性質進行不同遮罩方式
- 提供個資還原機制，以提升資料應用價值

資料表欄位設定遮罩

ArticleId	Category	Author
1	Development	xxxx
2	Testing	xxxx
3	DLM	xxxx
4	Development	xxxx
5	Testing	xxxx

Mask



ArticleId	Category	Author
1	Development	Atif
2	Testing	Peter
3	DLM	Sadaf
4	Development	Peter
5	Testing	Sadaf

UnMask

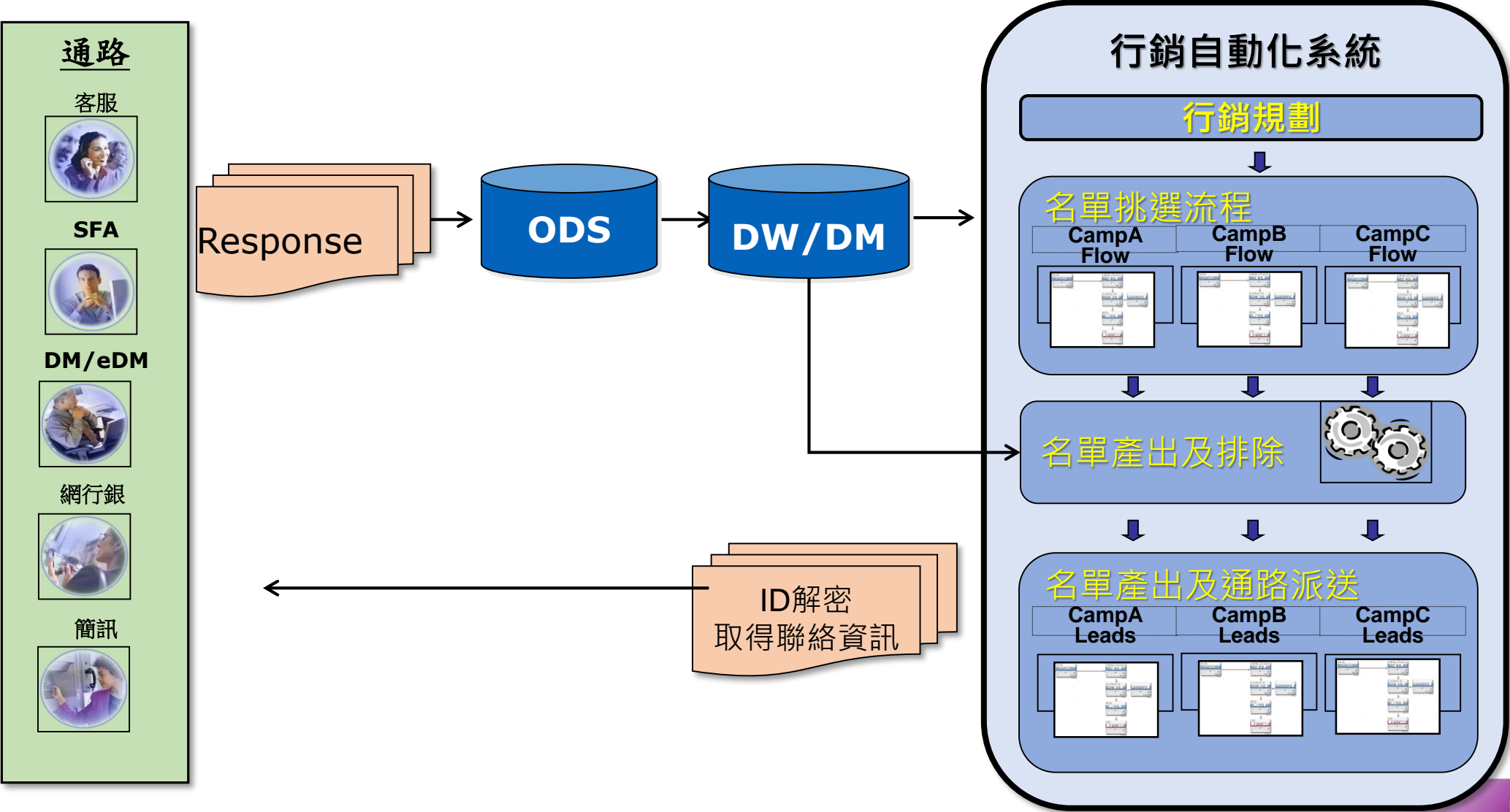


姓名	電話	地址
王小明	098765432	台北市松山區光復南路1號6樓
電子郵件信箱		
Ming.Wang@Company.com		



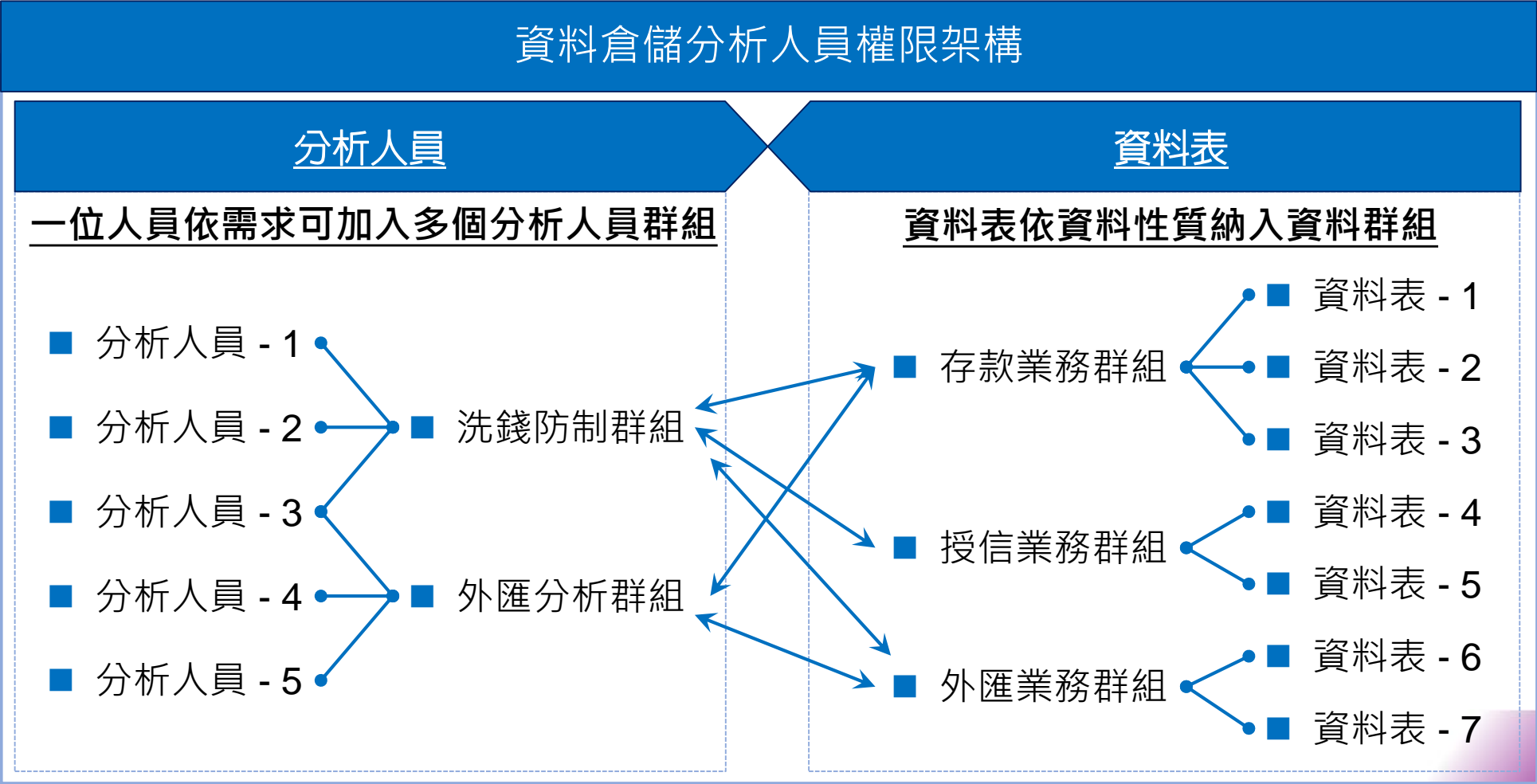
姓名	電話	地址
王 X X	0987*****	台北市松山區光復南路*****
電子郵件信箱		
xxxx@Company.com		

# 資料治理 – 行銷自動化系統應用案例



# 資料治理 – 分析人員權限管理架構

- 分析人員資料存取群組，依負責的業務或組織設定，做到最小授權
- 資料群組依資料業務性質分群

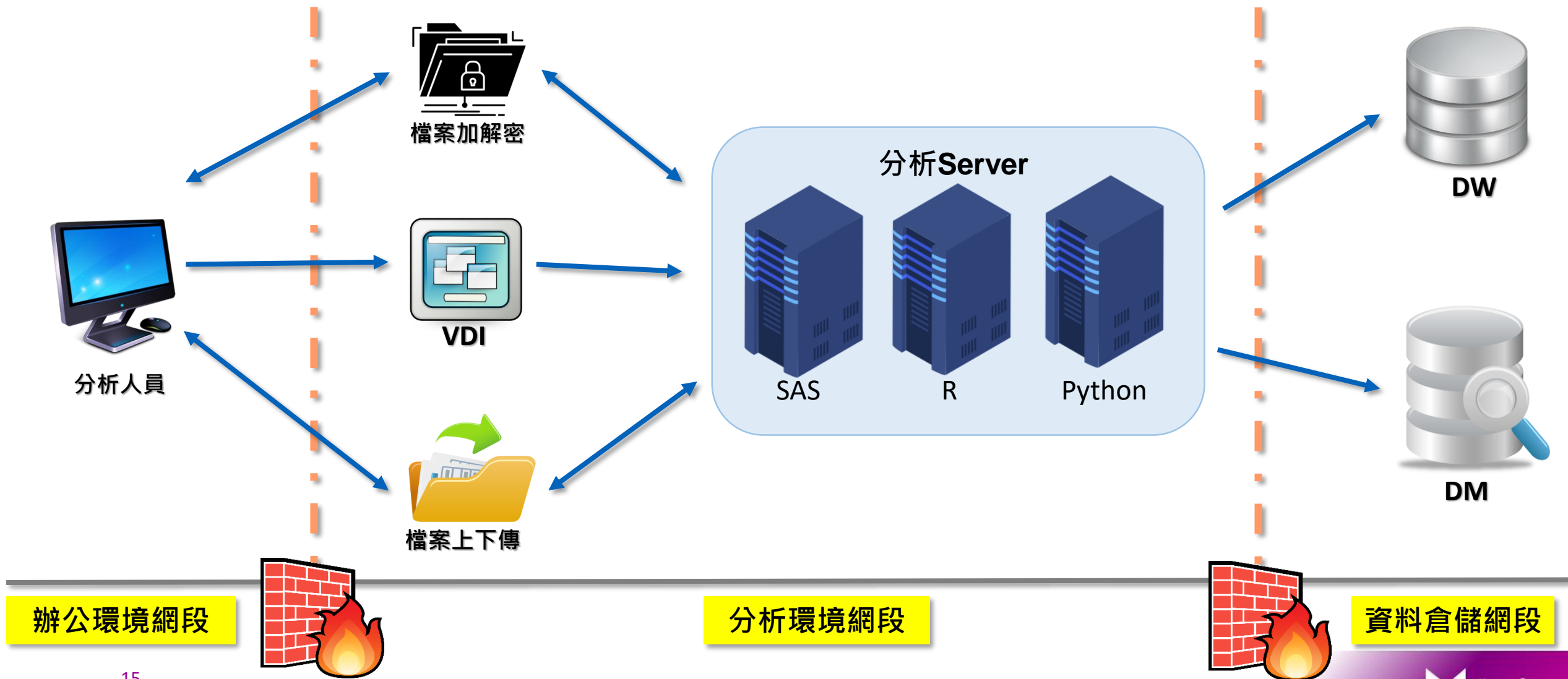


# 資料治理 – 群組權限關連表

■ 存取群組與人員之權限管理，設定於資料庫權限層級

		資料群組(Data Group)												
		DW資料									DM資料			
		行動銀行	推播系統	現金卡系統	內部息系統	債清系統	消金徵審	風險控管平台	代收付系統	...	CRM	PM	RM	CMS
使用者群組(User Group)	個金管理處	V	V	V	V	V	V				V	V		V
	財務管理處				V							V		
	數位金融處	V	V	V	V	V	V				V	V		V
	資訊處	V	V	V	V	V	V	V	V		V	V		V
	法金授信管理處								V					
	消費金融處	V	V	V			V				V			
	個金授信管理處					V	V	V			V			
	風險管理處							V					V	
	...													

# 資料治理 – 分析環境網路控管措施





# 資料治理 – 資料使用率分析

- 監控資料被使用狀況
- 運用DB 存取的Log機制產生制式報表
- 下架已無使用的資料，提昇空間效率
- 關閉已無登入的帳號，降低被駭風險

年月2009/04檢視報表

1之1100%尋找|下一個選取格式匯出

使用者使用狀況月報 2009/04

資料日期: 2009/04總頁數: 0001分頁數: 0001表號: DWM0001製表日期: 2009/04/20

使用者	資料超市系統名稱	使用次數	使用時間合計(秒)	使用時間百分比	CPU時間合計(秒)	CPU時間百分比	Physical I/O合計(Page)	Physical I/O百分比
crmuser	CRM資料超市	22	3.91	79.01 %	0.33	95.68 %	2,551	87.66 %
riskuser	信用風險資料超市	9	1.04	20.99 %	0.02	4.32 %	359	12.34 %

總辦覆核副科長科長製表 資訊室

年月2009/04檢視報表

1之1100%尋找|下一個選取格式匯出

資料表使用狀況月報 2009/04

資料日期: 2009/04總頁數: 0001分頁數: 0001表號: DWM0002製表日期: 2009/04/20

資料表名稱	使用次數	使用時間合計(秒)	使用時間百分比
CB_DT	20	2.10	42.41 %
CB_PCUST_MS	1	0.62	12.43 %
ODS_BD_BDWST	1	0.00	0.02 %
ODS_UN_CKMR	1	0.00	0.00 %
ODS_UN_FACT	1	0.00	0.00 %
ODS_UN_NLNR	2	0.79	15.94 %
ODSWS_UN_CFACT	1	0.00	0.00 %
ODSWS_UN_EDIACT	1	0.20	4.10 %
ODSWS_UN_NLNR	3	1.24	25.09 %

總辦覆核副科長科長製表 資訊室

年月2009/04使用者全部檢視報表

資料表名稱全部1之2100%尋找|下一個選取格式匯出

使用者與資料表使用狀況月報 2009/04

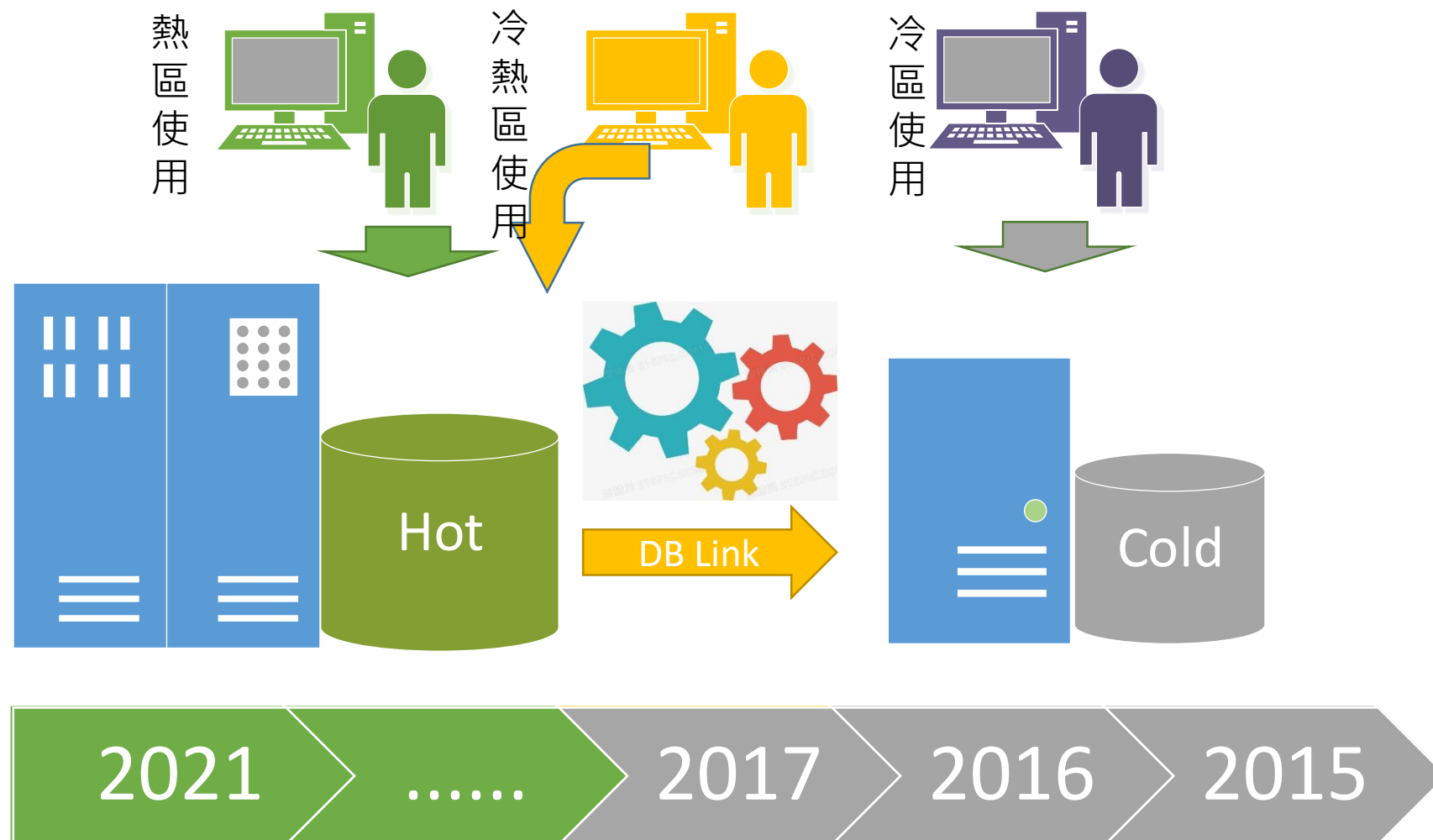
資料日期: 2009/04總頁數: 0002分頁數: 0001表號: DWM0004製表日期: 2009/04/20

使用者	資料超市系統名稱	資料表名稱	使用次數	使用時間合計(秒)	使用時間百分比	CPU時間合計(秒)	CPU時間百分比	Physical I/O合計(Page)	Physical I/O百分比
crmuser	CRM資料超市	CB_DT	20	2.10	42.41 %	0.21	59.87 %	1,615	55.50 %
crmuser	CRM資料超市	CB_PCUST_MS	1	0.62	12.43 %	0.09	27.09 %	58	1.99 %
crmuser	CRM資料超市	ODSWS_UN_NLNR	1	1.20	24.17 %	0.03	9.22 %	878	30.17 %
		小計	22	3.91	79.01 %	0.33	95.68 %	2,551	87.66 %
riskuser	信用風險資料超市	ODS_BD_BDWST	1	0.00	0.02 %	0.00	0.00 %	0	0.00 %
riskuser	信用風險資料超市	ODS_UN_CKMR	1	0.00	0.00 %	0.00	0.00 %	0	0.00 %
riskuser	信用風險資料超市	ODS_UN_FACT	1	0.00	0.00 %	0.00	0.00 %	0	0.00 %
riskuser	信用風險資料超市	ODS_UN_NLNR	2	0.79	15.94 %	0.02	4.32 %	64	2.20 %
riskuser	信用風險資料超市	ODSWS_UN_CFACT	1	0.00	0.00 %	0.00	0.00 %	0	0.00 %



# 資料治理 – 歷史資料管理與應用

- **熱區**存放常用分析需求資料，會使用高階磁碟，通常會保存3年內資料。
- **冷區**存放不常使用資料，使用低階磁碟，供臨時性歷史資料需求。
- 當資料超過期限，自動化搬移資料，將資料由熱區搬移至冷區上。
- 此機制在於提供**空間運用效率**、**縮短備份時間**、**節省硬體成本**。



# 資料治理 – 資料品質管理

## 品質管理確保資料的正確性

- Domain value 檢核分析
- 數值型、空值檢核分析
- 極大極小值分析
- 金額資料核帳(比對會計帳)檢核

資料品質分析 > 異常分析值維護 > 異常分析值查詢

資料庫名稱

資料表名稱

欄位名稱

查詢結果 共有 5 筆資料目前在第 1 頁，共有 1 頁。

資料庫名稱	資料表名稱	欄位名稱	檢核型態	合理範圍
ODSDB	ODS_CC_SCMCUSTP	MARITAL_ST	文字	0,1,2,3,4
ODSDB	ODS_CC_SCMCUSTP	QUALIFICATION	數值	0~9
ODSDB	ODS_CC_SCMCUSTP	SEX	文字	0,1
ODSDB	ODSMS_UN_PRSN	MARCD	文字	0,1,2,3
ODSDB	ODSMS_UN_PRSN	SEX	文字	0,1,2

資料品質分析 > 完整度分析值維護 > 完整度分析值查詢

資料表名稱

欄位名稱

查詢結果 共有 14 筆資料目前在第 1 頁，共有 1 頁。

資料庫名稱	資料表名稱	欄位名稱	警告百分比
ODSDB	ODS_CC_SCMCUSTP	ADDR1	90.00
ODSDB	ODS_CC_SCMCUSTP	DTE_BIRTH	90.00
ODSDB	ODS_CC_SCMCUSTP	MARITAL_ST	90.00
ODSDB	ODS_CC_SCMCUSTP	PHONE	60.00
ODSDB	ODS_CC_SCMCUSTP	QUALIFICATION	90.00
ODSDB	ODS_CC_SCMCUSTP	SEX	90.00
ODSDB	ODSMS_MF_CTMFIL	CTM09	90.00
ODSDB	ODSMS_MF_CTMFIL	CTM11	90.00
ODSDB	ODSMS_UN_PRSN	BIRTHDAY	90.00
ODSDB	ODSMS_UN_PRSN	JOB CD	90.00
ODSDB	ODSMS_UN_PRSN	LADDR1	90.00
ODSDB	ODSMS_UN_PRSN	MARCD	90.00
ODSDB	ODSMS_UN_PRSN	SEX	90.00
ODSDB	ODSMS_UN_PRSN	TELNO	90.00

資料輪廓分析 > 輪廓分析值維護 > 輪廓分析值查詢

資料庫名稱

資料表名稱

欄位名稱

查詢結果 共有 9 筆資料目前在第 1 頁，共有 1 頁。

資料庫名稱	資料表名稱	欄位名稱	資料分析型態
ODSDB	ODS_CC_SCMCUSTP	ADD_ZIPCODE	維度欄位
ODSDB	ODS_CC_SCMCUSTP	MARITAL_ST	維度欄位
ODSDB	ODS_CC_SCMCUSTP	QUALIFICATION	維度欄位
ODSDB	ODS_CC_SCMCUSTP	SEX	維度欄位
ODSDB	ODS_MF_CTMFIL	CTM12	維度欄位
ODSDB	ODSMS_UN_PRSN	JOB CD	維度欄位
ODSDB	ODSMS_UN_PRSN	LZIPCD	維度欄位
ODSDB	ODSMS_UN_PRSN	MARCD	維度欄位
ODSDB	ODSMS_UN_PRSN	SEX	維度欄位

# 資料治理 – MetaData管理

## ■ 資料定義說明管理，確保資料運用的正確性及提供分析人員訓練

- 資料欄位整理資料定義、商業邏輯相關MetaData
- 資料字典提供查詢，及訓練使用
- 業務規則整理，能確保資料定義的一致性

資料倉儲中介資料 > 欄位資訊檔明細

資料表格名稱：ODS\_BD\_BDCUST

查詢結果 共有33筆，共有1頁，目前在第1頁

資料欄位名稱	資料欄位中文名稱	主鍵記號	欄位型態
DATA DT	資料日期	Y	datetime
DEP_KEY	部門代號	Y	char(10)
CUST_NO	客戶代號	Y	char(12)
CUST_ID	統一編號	N	char(12)
CUST_FULL_NAME	客戶名稱	N	char(50)
CUST_SHORT_NAME	客戶簡稱	N	char(12)
CONTACTOR	連絡人(1)	N	char(14)
CUST_TEL_1	連絡人電話(1)	N	char(14)
FAX_NO_1	傳真號碼(1)	N	char(12)
CONTACTOR_2	連絡人(2)	N	char(14)
CUST_TEL_2	連絡人電話(2)	N	char(14)
CONTRACT_NO	契約編號	N	char(12)
CUST_RANKING	客戶評等	N	char(5)
CUST LOCATE CO	客戶歸屬	N	char(3)

資料倉儲中介資料 > 業務處理規則資訊 > 業務處理規則資訊查詢

☒ 輸入業務規則名稱

☐ 選擇業務規則名稱

☐ 輸入業務規則代號

應收利息

全部

☐ 完整比對

☒ 模糊比對

☐ 完整比對

☐ 模糊比對

搜尋

說明

查詢結果 共有 1筆資料目前在第 1 頁，共有 1 頁。

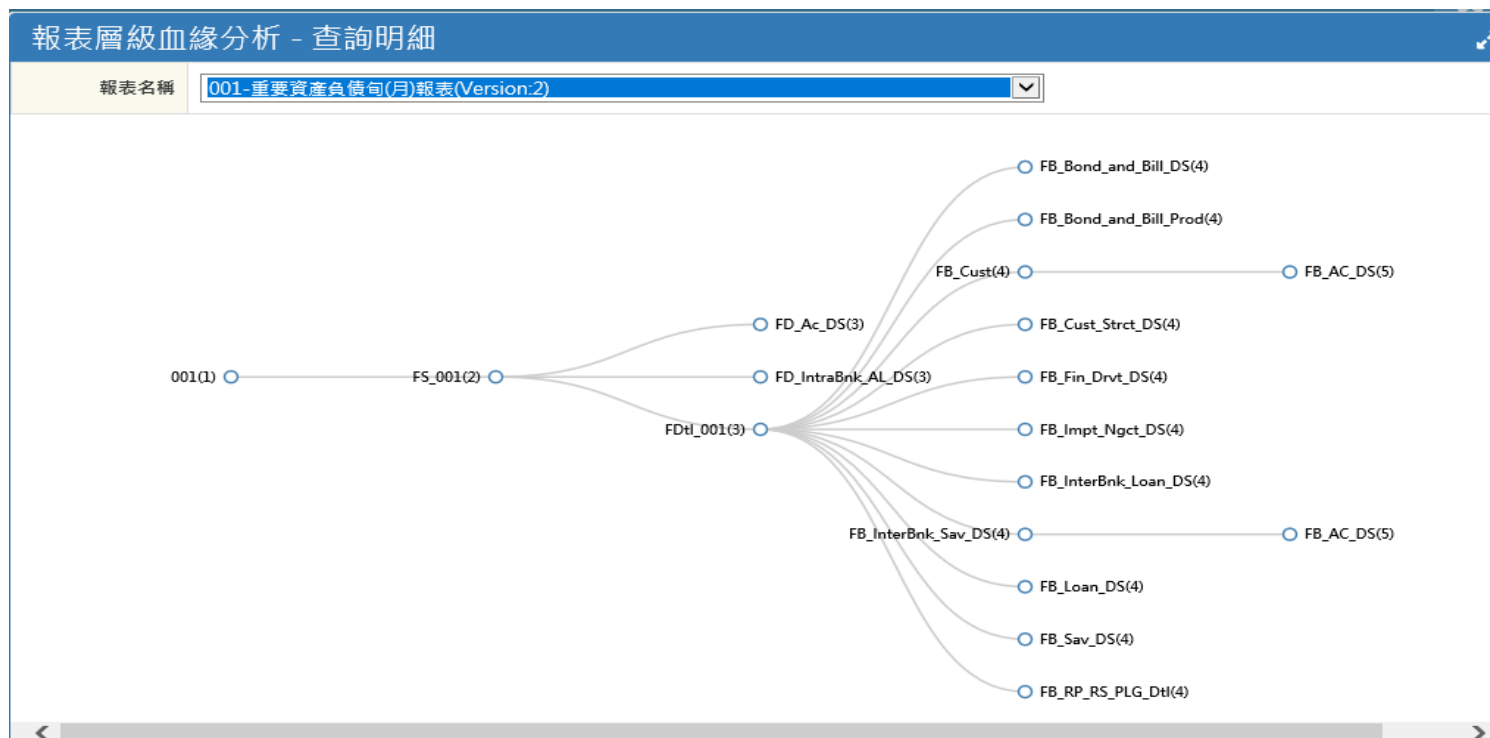
第一頁 上一頁 下一頁 最後一頁

業務規則代號	業務規則名稱	業務規則說明	連結Excel文件	維護人員編號
IntrIncomeAmt	當月應收利息	當月總積數 * 當月計息平均利率 / 365	開啟文件	026540

# 資料治理 – 資料流關連分析

## ■ 資料流程關連管理

- 上下游資料相依性關連資訊，提供資料異動分析
- 確保上游資料變更，下游系統的資料保持一致性與完整性



# 總結

- 資料倉儲業務領域中，數據分析便利與安全性並重。
- 資料治理使資料倉儲不會成為個資問題的根源。
- 適當的資料治理措施，有助於確保資料正確性、提昇作業流程效率、提高數據帶來的效益。



# Thanks