



# 資料治理—資料資產管理的最佳實踐框架

## 資料治理助力企業永續發展(1)

資料治理是為了確保權責及所有權而定義企業內資料資產的角色、責任和流程。本文讓你對資料治理有一個快速的理解。

文/Thor Olavsrud 譯/葉庭筠

在文章一開始，我們要先看看資料治理的定義。

資料治理是定義組織內誰對資料資產具有權威和控管權，以及如何使用這些資產的系統，包含了管理和保護資料資產所需的人、流程和技術。

資料治理研究院對資料治理的定義是「資訊相關流程的決策權利與責任系統，依大家同意的模型執行，這個模型描述在遇到何種情況下，透過何種方法，對哪些資訊採取何種行動。」

資料管理協會(Data Management Association, DAMA)將資料治理定義為「資料、資料使用及資料來源管理的規劃、監督和控管。」

## 資料治理 vs. 資料管理

資料治理是資料管理學科中的一部份，也是很重要的一部份。雖然資料治理談的是為確保權責與資料資產所有權的角色、責任和流程，不過DAMA對資料管理的定義是「一個總稱，描述用於規劃、指明、實現、建立、維護、使用、歸檔、取用、控管及清除資料的流程」。

資料管理成為很普遍的學科名詞，但有時它用於指稱資料資源管理或企業資訊管理(EIM)。Gartner指EIM是「架構、描述與治理跨組織和技術範疇的資訊資產，以提升效率、促進透明度和實現業務洞見的整合式學科。」

## 資料治理框架

資料治理可以想像成支援組織各種資料管理策略的功能。資料治理架提供貴公司一個從全觀式角度蒐集、管理、確保、儲存資料的框架。為協助企業了解資料治理框架應涵蓋什麼，DAMA將資料管理比喻做一個輪子，資料治理是軸承，周圍則為以下10個資料管理知識領域：

- 資料架構：資料和作為企業整體架構的相關資源組成的架構。
- 資料建模與設計：分析、設計、建造、測試和維護。
- 資料儲存和運作：結構化實體資料資產儲存的部署和管理。

- 資料安全：確保隱私、機密性及適切存取。
- 資料整合和相容性：資料取得、汲取、轉換、搬移、交付、複製、聯邦(federation)、視覺化及運算的支援。
- 文件和內容：儲存、防護、索引和提供存取非結構化資源內的資料，並使資料得以和結構化資料相容及整合。
- 資料倉儲和商業智慧(BI)：管理分析性資料的運算，並實現決策支援資料的存取，以便產出報表和分析。
- 參考和主資料(master data)：管理共享資料以減少重覆性，透過標準化定義及資料值的使用來提升資料品質。
- Metadata：主資料的蒐集、分類、維護、整合、控管、管理及交付。
- 資料品質：定義、監控、維護資料完整性及提升資料品質。

在建立資料治理策略時，上述每一個資料蒐集、管理、歸檔和使用各面向都需要考量到。

企業應用研究中心(Business Application Research Center, BARC)警告資料治理並不是一個「平地一聲雷」的單次計畫。作為一個長期性的複雜計畫，資料治理可能會發生參與成員漸漸失去信任或興趣的難題。為此BARC建議企業一開始推行可管理或應用專屬的試行專案，之後再把資料治理依據經驗推廣到全公司。

BARC建議以下策略來推行資料治理：

- 定義目標、了解好處。
- 分析現有狀態並做delta分析。
- 產生藍圖。
- 尋求股東或投資者的支持，找到預算。
- 發展和規劃資料治理方案。
- 實行資料治理方案。
- 監督和控管。

## 資料治理的目的

資料治理的目的是建立方法、權責及資料標準化、整合及儲存公司資料的流程。根據BARC指出，企業資料治理的目標應該要能：

- 將風險最小化。
- 建立資料使用的內部流程。
- 符合法遵要求。
- 強化內外部溝通。
- 提升資料價值。
- 促進管理。
- 降低成本。
- 透過風險管理和最佳化協助公司永續發展。

BARC指出，資料治理方案總是涵括企業的策略、戰術及作業層次，且必須是持續、反覆進行過程。

## 資料治理的好處

大部份企業其實已經為特定應用、業務單位或功能實行某些形式的資料治理，即使有些沒有正式的流程或責任。資料治理的目的在為這些流程和責任建立系統性的正式控管，這有助於協助企業維持靈活，特別是在它們逐漸茁壯到必須分工作業的規模時。資料管理作業的許多好處唯有企業建立起系統化的資料治理才會實現。其中好處包括：

- 統一而連貫的企業資料帶來更全面性、更好的決策資源。
- 流程和資料變更有清楚規範以協助提升企業和IT的敏捷和擴充性。
- 透過中央控管機制來降低其他資料管理領域的成本。
- 藉由重覆利用流程和資料提升效率。
- 改善資料品質及資料流程的文件化。
- 提升資料合規性。

## 資料治理原則

依據資料治理研究所，成功的資料治理和託管(stewardship)方案核心有8大原則：

1. 所有參與資料治理的成員彼此應誠實相待。他們在商討資料決策的驅動因素、限制、選項和影響時必須誠實而積極。
2. 資料治理和託管過程需要透明，需所有成員和稽核者都能明白如何在流程中導入資料決策和控管。

3. 資料治理的資料決策、流程和控管要被稽核，必須搭配文件化來支持合規及作業稽核要求。
4. 資料治理要定義誰負責跨功能的資料決策、流程和控管。
5. 資料治理需定義出誰要對資料託管者負責，後者可能是個別貢獻者或一組資料託管者(data stewards)的職責。
6. 制衡原則。資料治理要定義業務和技術部門之間，以及資料生產／蒐集、管理者、使用者以及標準與規範導入者的制衡權責。
7. 資料治理需導入和支援企業資料的標準化。
8. 變更管理。資料治理需支援參考資料值的主動和反應式變更管理，以及主資料及metadata的結構和使用。

## 資料治理角色

每個企業的資料治理方式都不相同，但還是有共通之處：

### 指導委員會

由於資料治理是涵括整個公司的，因此一般最先有一個指導委員會(Steering committee)，由高階主管組成（通常是CXO或部門副總裁）。《Get Governed: Building World Class Data Governance Programs》(www.amazon.com/Get-Governed-Building-Governance-Programs/dp/069295175X)一書作者Morgan Templar 指出，指導委員會成員的責任包括制訂整體治理策略和目標成果、支持資料託管的工作，並且要讓組織達成時程和成果。

### 資料持有人

Templar 表示，資料持有人(data owner)是負責確保特定資料領域的資訊獲得跨系統和部門的治理。它們通常是指導委員會成員，不過可能沒有投票權。資料持有人負責：

- 允許資料詞彙和其他資料定義。
- 確保全公司資訊的精準性。
- 指導資料品質活動。
- 審理和許可master data管理作法、成果和活動。
- 和其他資料持有人合作解決資料問題。

- 針對資料託管人找出的問題提供第二層審理。
- 針對其資料領域的軟體解決方案、政策或規範要求提供指導委員會意見。

### 資料託管人

資料託管人(data steward)負責資料的日常管理，Templar說，他們是特定領域的專家(subject matter experts, SME)，了解並和眾人溝通資訊的意義和使用，而且他們和公司內其他資料託管人一起扮演大部份資料決策的治理團體。資料託管者職責是：

- 是其資料領域的專家。
- 辨識出資料問題，並和其他資料託管人合作解決。
- 為資料託管委員會成員。
- 針對資料政策和委員會活動做出提案、討論與表決。
- 在資料領域人向資料持有人和其他利害關係者報告。
- 與各業務部門攜手，確保其領域資料獲得理解及適當管理。

### 資料治理工具

資料治理是一個持續性的過程而非單一技術方案，但是還是有工具可以支持它們。而什麼樣的工具適合貴公司則要視你的需求、資料量及預算而定。IT評等網站 IT Central Station ([www.itcentralstation.com/categories/data-governance](http://www.itcentralstation.com/categories/data-governance))指出，市場上頗受歡迎的資料治理解決方案有以下幾種：

#### SAS Data Management

在SAS平台上開發的SAS Data Management提供角色為基礎的流程管理GUI，整合了業務詞彙、SAS與第三方metadata管理、以及族系視覺化。

#### Collibra Governance

Collibra是全企業的資料治理解決方案，可自動化執行多種治理和託管任務，它內含政策管理員、資料服務台(helpdesk)、資料字典及業務詞彙。

#### Informatica Axon

Informatica Axon是一個支援資料治理方案的蒐集中心和資料交易市集。它主要功能包括協同式業務詞彙(glossary)、視覺化資料族系(lineage)及根據業

務定義來產出資料品質量測。

#### Varonis Data Governance Suite

Varonis解決方案可利用擴充式Metadata 框架自動化執行資料防護和管理任務，幫助公司管理資料存取、檢閱每個檔案、郵件事件的稽核紀錄、辨識不同業務單位的資料持有人，並且找出及分類敏感資料和文件。

#### Alation

Alation是企業自動化目錄，可依來源自動索引資料。它最主要功能TrustCheck可為workflow提供即時資料治理「護欄」。TrustCheck專為支援自助分析而設計，還可為資料資產加上指導方針及規則。

#### SAP Data Hub

SAP Data Hub是資料協作(orchestration)解決方案，旨在協助企業發現、精製、豐富化及治理各種類、各型態及大大小小的資料。它協助企業做好安全設定、為用戶、群組和角色找出控管政策，並且簡化最佳實踐和流程以進行政策管理及安全紀錄。

#### IBM Data Governance

IBM Data Governance運用機器學習來蒐集與策畫導引(curate)資料資產。它的整合式資料目錄協助企業尋找、策畫、分析、準備與分享資料。

#### Unifi Data Platform

Unifi Data Platform是一個自助式資料發現(data discovery)和準備的工具套件，它能自動化資料揣摩(data profiling)和加標籤、語義搜尋和自然語言查詢。其知識圖譜可從資料和來源之間的關係形成洞見。