

臺灣生物多樣性資訊平台 (TaiBIF) 建置建議

2004/12/08 MAAT

全球生物多樣性資訊中心 (Global Biodiversity Information Facility, GBIF) 是目前全球生物多樣性相關資訊網站或資料庫計畫中最重要, 層次最高的一個組織, 而臺灣目前係為此組織之附屬會員之一。應因 GBIF 對各參與國家、組織與團體提出的需求, 臺灣需設置一個國家節點 (national node) 作為統一對外聯絡之窗口。

臺灣生物多樣性資訊平台 (TaiBIF) 如果要成為這樣的一個國家節點, 它是應該要遵循 GBIF 所推薦使用的協定與資料傳輸格式, 如此上至 GBIF, 下至國內計畫, 才能表現整體貫徹。後設資料工作組根據 UDDI registry、DiGIR Protocol、Darwin Core Metadata Schema 等相關內涵習研與 GBIF 實例個案的觀察分析, 提出臺灣生物多樣性資訊平台 (TaiBIF) 建置建議。

(一) 臺灣生物多樣性資訊平台 (TaiBIF) 之建議架構

我們建議臺灣生物多樣性資訊平台 (TaiBIF) 之整體架構下圖所示。其構成元件, 主要有以下三個:

- Portal
- Registry (UDDI)
- Provider

元件之間除了必需依靠共同的通訊協定 (Protocol) 來完成溝通外, 也需一套共同的後設資料 schema 作為資料互通的機制。

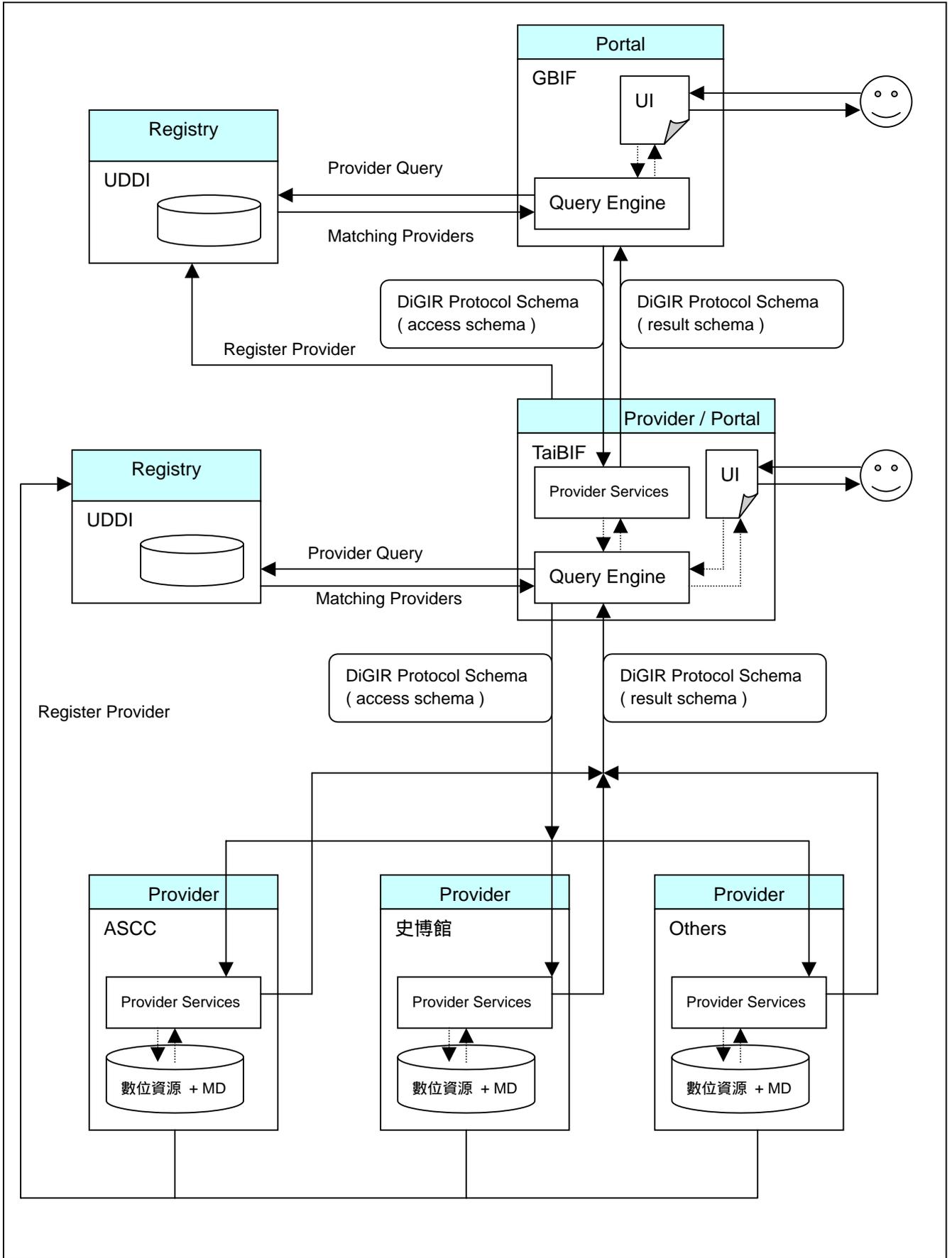
- DiGIR Protocol
- Metadata Schema (Darwin Core)

根據 DiGIR 網站公佈的兩份技術文件資料 (DiGIR Portal Reference Implementation Design : http://digir.sourceforge.net/port/portal_design.html ; Requirements and Assumptions : http://digir.sourceforge.net/gen/requirements_0_92.html), 各 DiGIR 元件、通訊協定與資料互通格式之所肩負的工作與功用性, 茲整理略述如下:

1. Portal

1. Portal 的部份至少需包含兩個部份, 分別為 UI (User Interface) 與搜尋引擎 (Query Engine)。
2. UI 需提供親善的使用者介面與 end user 相互溝通。
3. Query Engine 需以 DiGIR Protocol Schema 的方式傳送使用者之查詢條件至下層的 Provider 進行相關資料的查詢, 並需具備可接受 Provider 回傳 DiGIR Protocol Schema 結果之能力。經過回傳結果的剖析後, 呈現使用者查詢之最終結果。
4. 此平台需可連接到 Registry, 並可從中擷取所需之 Provider 資訊。

5. 平台可定時發出檢查 (checks), 測試 Provider 當時的狀況。
 6. 平台需有能力處理 Provider 回傳的錯誤情況 (error conditions)。
 7. 當查詢結果時間過長, 平台有能力終結該項查詢動作。
2. Registry (UDDI)
 1. 提供註冊 (registration) 的機制, 讓符合規定的平台註冊成為 Portal 下之 Provider。
 2. 需有能力貯存 Provider 的名字、存取、服務等相關的資訊。
 3. 可回覆 Portal 查詢有關 Provider 的資訊。
3. Provider
 1. Provider 需提供 Provider Services, 可接受 Portal 提出 DiGIR Protocol Schema 的查詢, 並回覆 DiGIR Protocol Schema 的查詢結果。
 2. Provider 可提供 Metadata 紀錄的 XML 格式文件, 並提供該 XML 格式文件對應的數位物件。
 3. Provider 在提供服務前, 需先到 Registry 進行註冊的動作。
 4. 系統需維護 metadata 的相關資訊, 並對這些資訊進行郵戳標示, 當資訊有更新時, 即可立即通知 Portal 進行相對應的更新。
 5. 當 Provider 無法完成 Portal 提出的查詢時, 需回應 Portal 錯誤的訊息。
 6. 可接受來自 Portal 定期的查檢 (checks) 訊息, 並回覆 Portal 目前 Provider 的運作是否一切正常。
4. DiGIR Protocol
 1. DiGIR Protocol 需定義擷取 (retrieve) metadata 資訊之格式。
 2. DiGIR Protocol 需定義回應 (return) metadata 資訊之格式。
 3. DiGIR Protocol 需定義表達錯誤的格式與錯誤碼。
 4. DiGIR Protocol 需定義 Portal 定期查檢訊息的格式。
5. Metadata Schema (Darwin Core)
 1. 以 Darwin Core 作為 Portal 與 Provider 間資料互通使用之 metadata schema。
 2. 各 Provider 上傳相同格式 (schema) 之 metadata 到 Portal 去, Portal 進行整合後以相同的格式呈現讓使用者查詢與瀏覽。



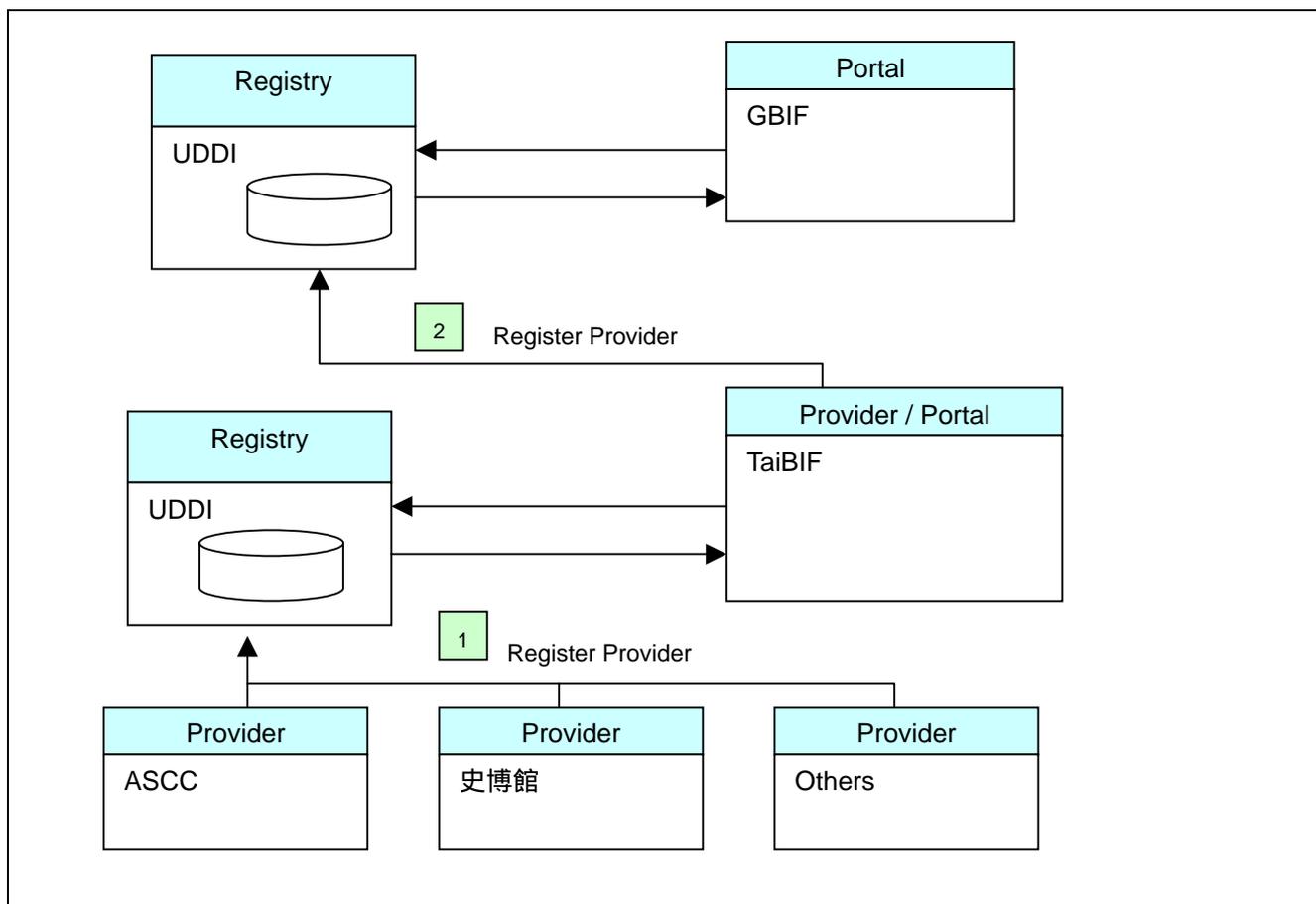
圖一：臺灣生物多樣性資訊平台 (TaiBIF) 元件架構圖

(二) GBIF 與 TaiBIF 相互運作流程

臺灣生物多樣性資訊平台 (TaiBIF) 正式成為 GBIF 之臺灣地區節點，讓國際使用者透過 GBIF 檢索到台灣地區生物多樣性資訊前，TaiBIF 平台需先完成 TaiBIF UDDI Registry 元件的建置與 GBIF UDDI Registry 註冊作業：

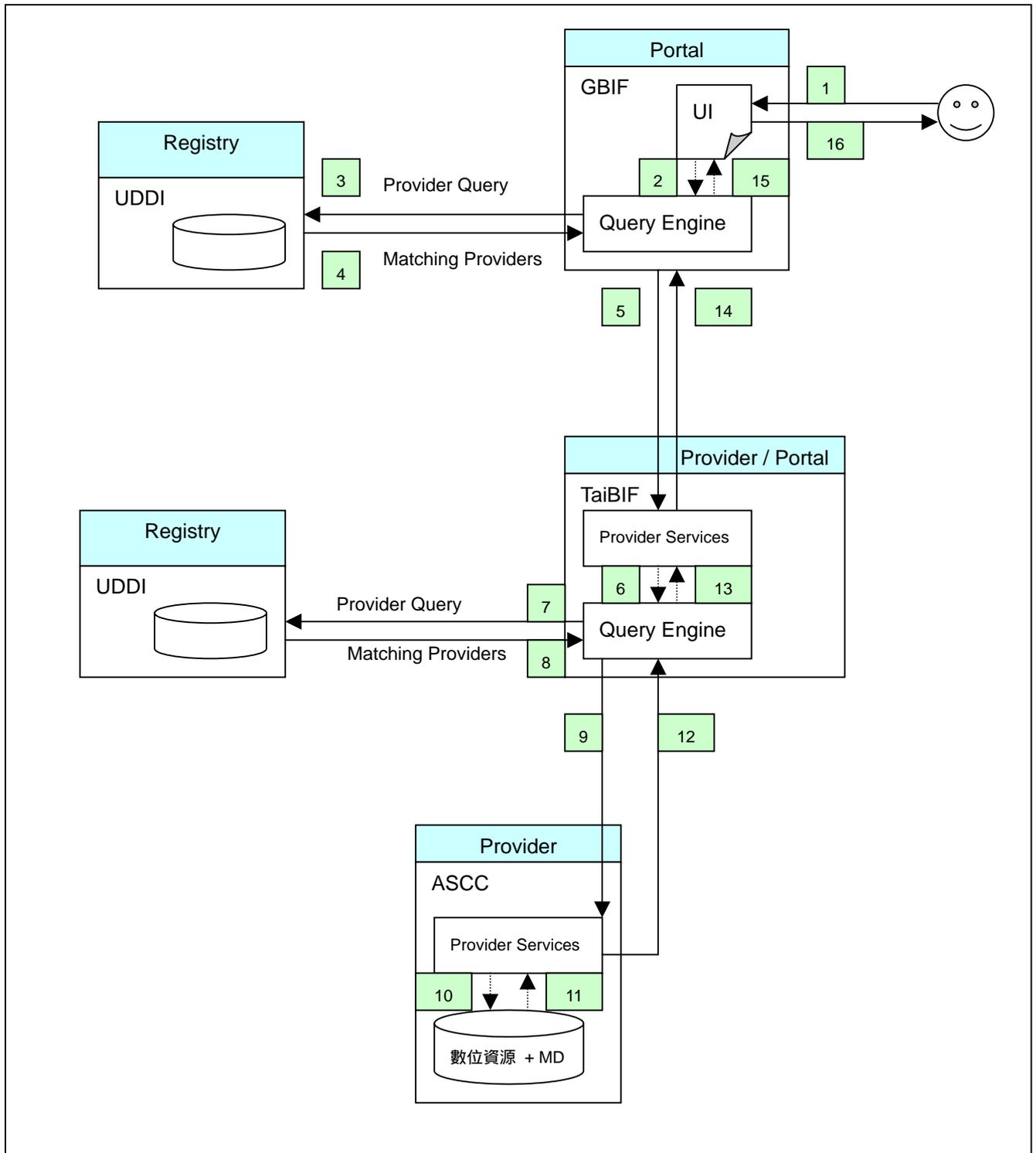
1. TaiBIF UDDI Registry 建置：
TaiBIF 系統建置人員需先建置 TaiBIF 平台與其 UDDI Registry，並提供參與生物多樣性之國內計畫 (Provider) 註冊其相關的資訊，如：實體名稱、實體入口點位址、服務名稱與類型等。
2. GBIF UDDI Registry 註冊：
TaiBIF 平台之 UDDI Registry 接受完各計畫 (Provider) 的註冊後，TaiBIF 平台即可向 GBIF 平台之 UDDI Registry 註冊成為 GBIF 下臺灣地點之節點 (national node)，並把各計畫註冊在 TaiBIF UDDI Registry 的資訊依 UDDI 規定之資料模式，如：BusinessEntity、BusinessService、BindingTemplate、tModel 與 PublisherAssertion 等，提供給 GBIF UDDI Registry，如此當國際使用者欲檢索臺灣地區之生物多樣性資料時，GBIF 平台才有足夠的資訊與 TaiBIF 平台溝通，並進一步取得檢索結果回覆使用者。

有關 TaiBIF 與 GBIF 之 UDDI Registry 註冊示意圖如下所示：



圖二：TaiBIF 與 GBIF 之 UDDI Registry 註冊示意圖

當 TaiBIF 完成註冊後，國際使用者即可對 TaiBIF 下之 Provider 作出生物資訊的搜尋，以下是搜尋動作之運作流程圖：



圖三：透過 GBIF 搜尋臺灣 Provider 生物資料運作流程示意圖

上圖詳細的運作流程說明茲分述如下：

步驟 1：

國際的使用者透過 GBIF 網站提供的使用者界面 (User Interface) 下達查詢條件，並指定欲查詢的 Provider Name (如：ASCC)。

步驟 2：

此查詢訊息隨即被傳送到 GBIF 之 Query Engine 去進行相關剖析。

步驟 3、4：

Query Engine 剖析完後，即連接到 GBIF UDDI Registry 去查詢有關 Provider 的資訊，如：使用者指定的 Provider 屬於哪個國家節點 (如：TaiBIF)？這個國家節點實際位址在哪？

步驟 5：

擷取足夠的資訊後，GBIF Query Engine 即可連線到該 Provider (ASCC) 的國家節點 (TaiBIF)，並把使用者下達的查詢條件轉成 DiGIR Protocol Schema 形式傳送到使用者欲查詢 Provider 隸屬的國家節點去。

步驟 6：

國家節點 (TaiBIF) 的 Provider Services 元件負責接受 GBIF Query Engine 傳來的 DiGIR Protocol Schema 形式查詢資訊，隨即將此訊息轉送到國家節點 (TaiBIF) 之 Query Engine，由 Query Engine 進行分析並傳送出去。

步驟 7、8：

國家節點 (TaiBIF) 之 Query Engine 分析完後，需到自己的 UDDI Registry 去詢問使用者欲查詢的 Provider (ASCC)，它實際位置在哪。

步驟 9：

獲知 Provider (ASCC) 實際位置後，Query Engine 即可把查詢訊息 forward 給 Provider。

步驟 10、11：

Provider (ASCC) 接受到來自國家節點 (TaiBIF) DiGIR Protocol Schema 格式的查詢後，就可以到自己的 Database 檢索有沒有符合使用者查詢條件的資料。

步驟 12：

Provider Service 元件會把查詢結果轉換成 Darwin Core 格式的 XML 文件，並依循 DiGIR Protocol Schema 形式把結果回傳到國家節點 (TaiBIF)。

步驟 13、14：

國家節點 (TaiBIF) 直接把此結果 forward 給 GBIF。

步驟 15、16：

GBIF 之 Query Engine 接受到 Provider 回傳之查詢結果後，進行過濾、剖析與重新包裝，最後藉由 User Interface 把使用者檢索到的結果呈現給使用者。

(三) 臺灣生物多樣性資訊平台 (TaiBIF) 建置工作項目與工作分配建議

一、 工作項目與工作分配建議

為了方便討論，MAAT 先將整個環境切割成以下幾個部份：

- GBIF 部份：包括的元件有 GBIF Portal 與其 UDDI Registry。
- TaiBIF 部份：包括 TaiBIF 平台與 TaiBIF UDDI Registry。
- Provider 部份：各參與計畫之 Provider。

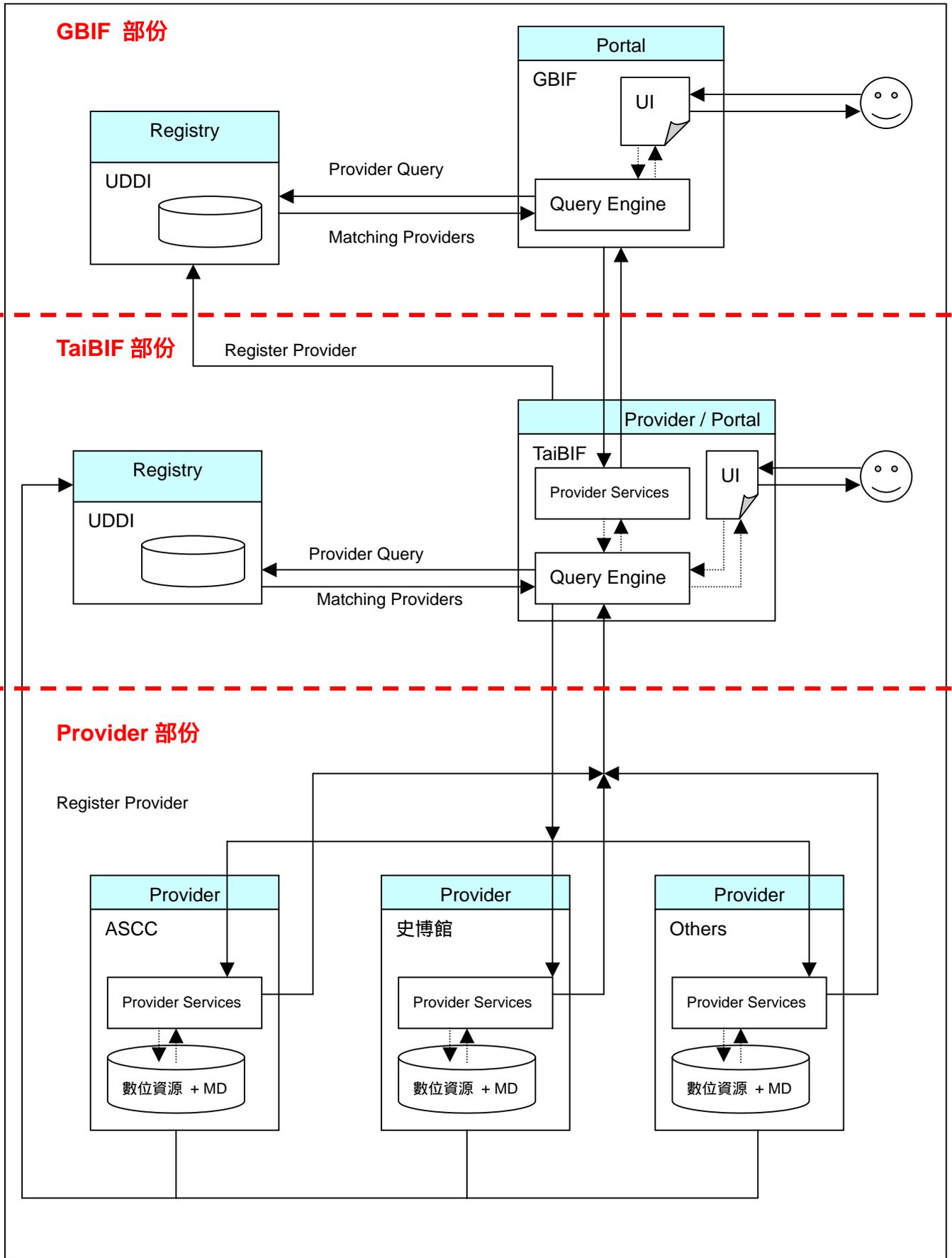
(有關範例的劃分，請見圖四：臺灣生物多樣性資訊平台工作範圍劃分)

然後細列上述各部份所該執行的工作項目與分配建議：

(表示負責、參與該項作業之單位。)

工作項目	負責單位		
	主題計畫	TaiBIF 系統建置	MAAT
前置作業			
了解此計畫相關的背景，並思考參與此計畫的目的，與在計畫中之定位、扮演的角色。			
研習相關的內涵。	若 provider 有足夠的人力或資訊人員，建議也可研習相關內容。	如 UDDI Registry、DiGIR Protocol、Darwin Core 等。	如 UDDI Registry、DiGIR Protocol、ABCD Schema、Darwin Core 等。
進行 GBIF 提供之 Toolkit 測試與評估。			
進行 Darwin Core 之中文化，與訂定 Darwin Core base 之 TaiBIF 版本與相關的 XML DTD，作為暨符合 GBIF 規定的資料描述格式，也符合臺灣生物多樣性描述格式的需求。			
GBIF 部份			
GBIF 部份的工作項目由 GBIF 建置人員處理，主題計畫、TaiBIF 系統建置者與 MAAT 暫無實際工作項目與分工。			
TaiBIF 部份			
建置 TaiBIF 平台，包括使用者介面設計。			
建置 TaiBIF UDDI Registry。			
測試 TaiBIF 與 GBIF 的連線，確保雙方可以相互傳達正確的資訊。			
測試 TaiBIF 與 UDDI Registry 的連線，確保兩者可以相互傳達正確資訊。			
測試 TaiBIF 與 Provider 的連線，確保雙方可以相互傳達正確的資訊。			
到 GBIF UDDI Registry 進行註冊。			
發展 Provider 工具，協助各主題計畫安裝或開發 Provider。			
TaiBIF 系統與 TaiBIF UDDI Registry 之維護。			
Provider 部份			

工作項目	負責單位		
	主題計畫	TaiBIF 系統建置	MAAT
安裝/開發 DiGIR Provider 的機制。		應事先制定或提供工作手冊或說明手冊，說明安裝的方式、步驟等等資訊，提供給 Provider。	
制訂資料庫欄位與 Darwin Core 之 mapping 關係。			MAAT 協助進行檢查的工作。
設定 DiGIR Provider 機制與資料庫的連接，並確保資料庫原資料可順利轉成 Darwin Core 格式之 XML 文件。			
向 TaiBIF UDDI Registry 進行註冊。			
測試 TaiBIF 與 Provider 的連線，確保雙方可以相互傳達正確的資訊。			
檢查資料庫原資料的資料完整性，並核訂是否有些資料不宜提供。			
Provider 系統的維護。			
後續作業			
進行整體的測試。			



圖四：臺灣生物多樣性資訊平台工作範圍劃分

(四) 參考資料

1. GBIF website : <http://www.gbif.org/>
2. TaiBIF website : <http://gbif.sinica.edu.tw/home.htm>
3. DiGIR website : <http://digir.sourceforge.net/>
4. UDDI website : <http://uddi.org>
5. UDDI Spec Technical Committee Specification, UDDI version 3.0.1
6. Understanding Web Service-XML, WSDL, SOAP, UDDI ; Eric Newcomer