



Riassunto della LERU roadmap per dati della ricerca

La tabella di marcia (roadmap) LERU per dati della ricerca ([LERU Roadmap for Research Data](#)) indica un percorso che la Lega Europea delle Università di Ricerca ([LERU](#) - League of European Research Universities), **e rispettivamente qualsiasi altra organizzazione di ricerca**, potrà percorrere a propria discrezione per implementare delle solide pratiche di gestione dei dati della ricerca a livello istituzionale. La tabella di marcia (roadmap) è suddivisa in sei capitoli, integrati dal capitolo sette che comprende una lista di raccomandazioni per i vari gruppi destinatari.

Il **capitolo 1°** contempla i concetti di **politiche e leadership (Policy and Leadership)** illustrando come le università e le organizzazioni di ricerca abbiano dato seguito in modo più o meno attivo alle direttive riguardanti le politiche circa il trattamento dei dati. Inoltre si afferma che per una valida gestione dei dati della ricerca sarebbe necessario introdurre delle politiche istituzionali per la gestione dei dati che andrebbero accompagnate da tabelle di marcia strategiche che ne illustrino la gestione.

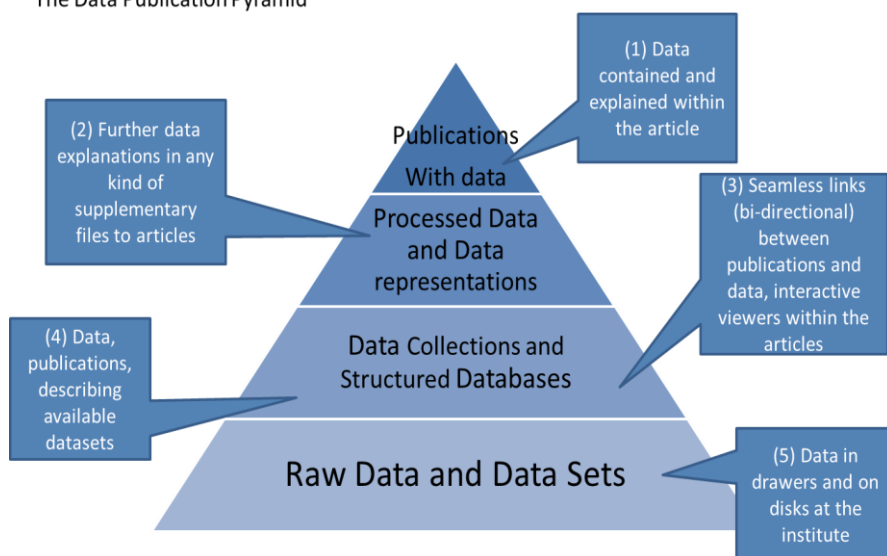
Il **capitolo 2°** tratta l'argomento del **patrocinio (Advocacy)** che secondo la tabella di marcia è d'importanza cruciale per una riuscita condivisione dei dati (Data Sharing). La tabella di marcia identifica sia incentivi sia ostacoli per una condivisione dei dati, indicando anche proposte su come si potrebbero superare eventuali riluttanze dei ricercatori rispetto al rilascio dei propri dati.

Uno degli obiettivi da raggiungere è la produzione di dati della ricerca aperti, e ciò - laddove possibile - dovrebbe valere per tutti i ricercatori. Per questo si richiede un'adeguata leadership a livello istituzionale. Le università e i servizi di supporto delle organizzazioni di ricerca sarebbero enti particolarmente adatti a patrocinare le migliori pratiche nella gestione dei dati della ricerca e nella citazione di dati. Grazie a questo patrocinio si potrà mostrare l'utilità inerente alla condivisione dei dati che allo stesso tempo contribuisce a renderli visibili, a incrementarne il riutilizzo e a promuovere la collaborazione. In questo modo sarebbe possibile instaurare la fiducia necessaria a far sì che tutto ciò si realizzi.

Il **capitolo 3°** tratta di tutta una serie di temi che rientrano nell'ambito della **gestione dei dati della ricerca**

(Management of Research Data): selezione e raccolta (Selection and Collection), cura (Curation), descrizione dei dati (Description), citazione e questioni giuridiche (Citation and Legal Issues). Per la selezione e l'aggiornamento dei dati, la tabella di marcia fa riferimento alla ODE Data Publication Pyramid ¹ e raccomanda che le comunità dei ricercatori continuino ad impegnarsi nell'identificare quali strati della piramide potranno essere resi accessibili per una condivisione e un riutilizzo e quali potranno essere aperti. Per quanto riguarda la cura dei dati, la tabella di marcia analizza i

The Data Publication Pyramid



¹ http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/11/ODE-ReportOnIntegrationOfDataAndPublications-1_1.pdf



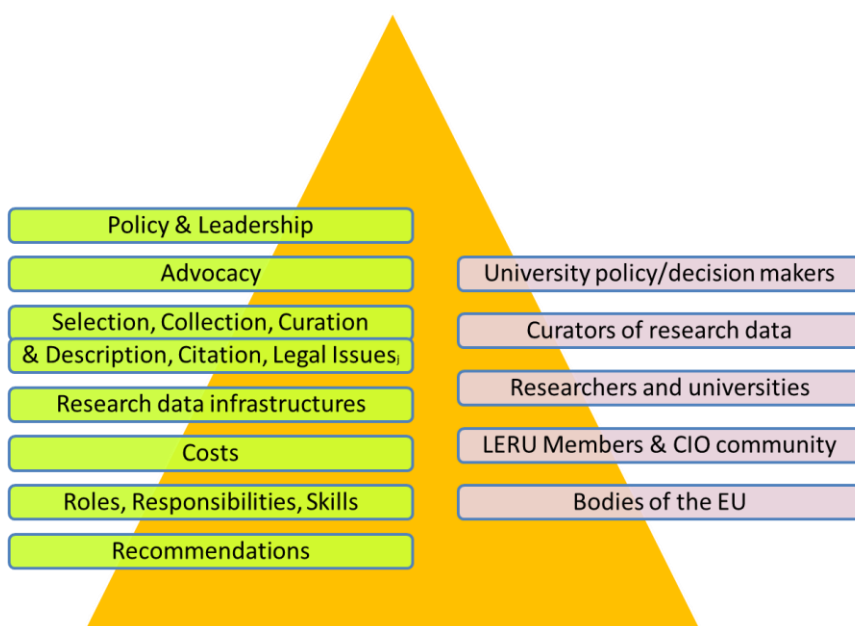
processi di ricerca suggerendo in seguito come si potrebbero creare le infrastrutture necessarie per sostenere questi processi. Per quanto riguarda la descrizione dei dati, la tabella di marcia fa notare le difficoltà inerenti al tentativo di incoraggiare i ricercatori a descrivere in modo dettagliato i propri dati. Nella sezione dedicata alle citazioni, vengono elencati esempi di buona pratica. La parte riguardante le questioni giuridiche analizza le condizioni quadro del copyright a livello europeo e suggerisce che sarà necessaria una deroga per il testo e l'estrazione di dati (Data Mining) per favorire uno sviluppo positivo dell'estrazione del contenuto (Content Mining) in Europa.

Il **capitolo 4°** è dedicato all'**infrastruttura dei dati della ricerca (Research Data infrastructure)**. Queste infrastrutture possono essere suddivise in quattro tipi:

- dati della ricerca in senso proprio
- gestione dei dati
- tool/strumenti per la gestione di dati
- aspetti tecnici, reclutamento del personale

L'infrastruttura dei dati di ricerca deve offrire delle condizioni quadro generiche che tengano conto della grande varietà delle attività di ricerca che le utilizzeranno. In questa sede viene offerta una visione panoramica dei tool della gestione dei dati della ricerca e il capitolo fa riferimento alla cosiddetta "lunga coda dei dati", cioè a tutti quei dati che si trovano immagazzinati su desktop locali, hard disk e server, e che probabilmente costituiscono una sfida maggiore rispetto ai cosiddetti "big data". Per quanto riguarda gli aspetti tecnici, il capitolo illustra questo mondo trasversale di aspetti tecnici presenti all'interno di tutta l'istituzione. Si conclude suggerendo che, preferibilmente, i servizi di supporto dovrebbero essere organizzati come unità omogenea.

Il **capitolo 5°** affronta il tema complesso dei **costi**. Non esiste neanche un singolo modello che potrebbe essere utilizzato per il calcolo dei costi. Questo capitolo offre due studi di casi specifici, e precisamente per l'Università di Oxford e per l'UCL (University College London), che indicano dei costi approssimativi per l'erogazione del servizio. Il capitolo mostra come l'analisi costi-benefici a volte faciliti la pianificazione dei costi e serva a valutare l'effettività della cura dei dati di ricerca in termini di costi. Viene anche indicato chi presumibilmente si farà carico dei costi – il finanziatore della ricerca, un servizio collaborativo nazionale o anche le università oppure le organizzazioni di ricerca per conto proprio.



Il **capitolo 6°** è dedicato a **ruoli, responsabilità e abilità (Roles, Responsibilities and Skills)**. Questo capitolo esegue un'analisi dei vari ruoli che sono necessari per la gestione dei dati della ricerca e rispettivamente coinvolti in questo processo e delle responsabilità in capo ai rispettivi collaboratori. Il capitolo suggerisce che la nuova figura professionale del Data Scientist possiede il potenziale per assumere un ruolo nuovo, del tutto indipendente. Il capitolo identifica anche le esigenze di formazione che dovranno essere soddisfatte per tutta una serie di

partecipanti a questi processi, p. es. studenti di formazione post-laurea/dottorandi o anche ricercatori senior, bibliotecari e data scientists.



Il **capitolo 7°**, quello finale, presenta **44 raccomandazioni** tratte dalla tabella di marcia e attribuite agli specifici gruppi target: gli operatori dei processi decisionali a livello istituzionale, tutte le figure coinvolte nella raccolta dei dati della ricerca, i ricercatori e le loro istituzioni, i membri di LERU e della comunità LERU dei Chief Information Officers (CIO, ossia i responsabili dei dipartimenti ICT) nonché gli enti dell'Unione Europea.

Paul Ayris, UCL

