

Strategia per la valorizzazione del patrimonio informativo del CRA

Premessa

L'art. 9 del DL n. 179/2012, convertito in Legge n. 221/2012, stabilisce di pubblicare nella sezione Amministrazione trasparente del sito web dell'Ente *“il **catalogo dei dati**, dei **metadati**, e delle relative **banche dati** ... ed i regolamenti che ne disciplinano l'esercizio della facoltà di accesso telematico e il riutilizzo ...”*. Lo stesso articolo stabilisce che i dati di tipo aperto (*Open Data*) sono *“i dati che ... sono disponibili secondo i termini di una licenza che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato ...”*, introducendo il principio dell'*Open by default*.

Con *“open data”* si indicano quelle tipologie di dati liberamente disponibili a tutti, privi di *copyright* o altre forme di controllo che ne impediscono la riproduzione, e le cui restrizioni si limitano all'**obbligo di citare la fonte**. Il movimento *Open Data* si richiama agli altri movimenti *“Open”*, come *Open Source*, *Open Access* e *Open Government*. Aderire al movimento *Open Data* significa iniziare a pubblicare sul *Web* i propri *dataset*, in un formato aperto, adatto ad essere impiegato a prescindere dagli strumenti informatici necessari per il loro successivo trattamento.

A conferma dell'importanza degli *Open Data* come *driver* di nuove forme di economia, la Commissione Europea ha inserito il tema nella *Digital Agenda - Pillar I: Digital Single Market, Action 3: Open up public data resources for re-use*¹.

Inoltre, nel *Summit G-8* del 2012 a Camp David è stato preso l'impegno di istituire la *“G-8 International Conference on Open Data for Agriculture”*, la cui prima edizione si è tenuta a Washington il 29-30 Aprile 2013. L'Italia è stata rappresentata dall'Ambasciata italiana, coadiuvato scientificamente dal CRA per conto del MIPAAF.

In quella sede l'Italia ha assunto l'impegno di aderire al movimento e ha presentato il proprio *Action Plan*² da attuare secondo quanto previsto nel *“Charter on open data signed by G8 leaders to promote transparency, innovation and accountability”*³.

L'impegno preso dall'Italia in sede G-8 si riversa sul Ministero delle politiche agricole alimentari e forestali, e su tutte le istituzioni italiane che operano nel campo dell'Agricoltura, tra cui il CRA.

Consapevole di possedere un patrimonio informativo di grande pregio, e di poter fornire in modalità aperta dati potenzialmente di grande valore, il CRA aderisce quindi al movimento *Open Data* al fine di poter contribuire allo sviluppo di un'economia basata sulla conoscenza e sull'innovazione.

¹ European Commission, Digital Agenda for Europe, *Pillar I: Digital Single Market, Action 3: Open up public data resources for re-use* <http://ec.europa.eu/digital-agenda/en/pillar-i-digital-single-market/action-3-open-public-data-resources-re-use>

² *G8 International Conference on Open Data for Agriculture*, Press Statement and Action Plans: <http://sites.google.com/site/q8opendataconference/action-plans>

³ *Charter on open data signed by G8 leaders to promote transparency, innovation and accountability*, Organisation: Cabinet Office, Page history: Published 18 June 2013 Policy: Improving the transparency and accountability of government and its services; Topical event: UK Presidency of G8 2013: <https://www.gov.uk/government/publications/open-data-charter>

Informazione scientifica e *Data Economy*

Lo sviluppo di un'economia basata sulla conoscenza e sull'innovazione è una delle priorità della strategia Europa 2020. Per questo scopo è considerato essenziale dare ampia diffusione alla ricerca finanziata con fondi pubblici, mediante la pubblicazione di dati e studi scientifici ad accesso aperto. Nella Raccomandazione della Commissione del 17 luglio 2012 sull'accesso all'informazione scientifica e sulla sua conservazione, si raccomanda agli stati membri di dare "accesso aperto alle pubblicazioni scientifiche, ai dati della ricerca e alla conservazione e riutilizzo dell'informazione scientifica"^{4, 5}.

Per "informazione" si intende tutto il patrimonio di conoscenza formalizzato attraverso l'uso di una grammatica di un qualsiasi linguaggio naturale (testo), ovvero attraverso notazioni matematiche, logiche, ecc., ma anche attraverso immagini, grafici, audio, video e quant'altro possa essere recepito dall'intelletto e dai sensi umani. Nel contesto di questo documento, con "informazione" si intende quindi una qualsiasi conoscenza contenuta all'interno di un documento, sia esso su supporto fisico (p.e. carta) che dematerializzato (p.e. in formato elettronico).

Con "dato" si intende qualunque informazione rappresentata in modo da poter essere trattata da un calcolatore elettronico. In questo senso, il formato sorgente di una pagina *Web* è un insieme di dati utilizzato da un particolare *software*, detto *browser* (p.e. Explorer). Viceversa, la pagina *Web* mostrata dal *browser* è informazione, poiché i dati sono mostrati dal *computer* in una forma adatta a trasferire conoscenza ad un essere umano.

Un esempio di dati particolarmente significativi per il CRA è rappresentato dai dati del Registro Nazionale delle Varietà di Vite⁶ realizzato dal Servizio Nazionale di Certificazione della Vite operante presso CRA-VIT di Conegliano. I dati sono mantenuti all'interno di una Banca Dati che contiene inoltre ulteriore informazione tra cui fotografie (*photo gallery*), grafici (serie storiche), mappe, ecc. La Banca Dati è corredata di un'apposita applicazione *software* che, oltre a permettere la gestione dei dati, consente di dare accesso all'informazione attraverso Internet per mezzo di un *browser*. Ovviamente possono esistere Banche Dati non accessibili dalla Internet. Ciò non toglie che i dati contenuti in queste Banche Dati non possano essere di grande interesse e rilevanza scientifica.

L'informazione sull'esistenza di una Banca Dati e sulla tipologia dei dati contenuti in essa, risulta quindi di enorme importanza per gli scopi dell'azione di valorizzazione del patrimonio informativo.

Una collezione omogenea di dati (*dataset*) è considerabile alla stregua di una più complessa Banca Dati, poiché tecnicamente è sempre possibile ridurre una Banca Dati in una collezione di *dataset*. Perciò, un *dataset* contenente dati rilevanti ma non disponibili attraverso una specifica applicazione *software* (p.e. disponibili solo come file

⁴ Gazzetta ufficiale dell'Unione europea - L 194/39 Raccomandazione della Commissione del 17 luglio 2012 sull'accesso all'informazione scientifica e sulla sua conservazione <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:IT:PDF>

⁵ A conferma di questo, Neelie Kroes, Vice-Presidente della CE, responsabile della Digital Agenda, in un recente discorso tenuto in occasione del lancio della Research Data Alliance ha affermato: "we are putting openness at the heart of EU research and innovation funding" (Opening up Scientific Data, Launch of the Research Data Alliance, Stockholm, 18/03/2013 European Commission - SPEECH/13/236 <http://ec.europa.eu/digital-agenda/en/news/opening-scientific-data>

⁶ Registro Nazionale delle Varietà di Vite <http://catalogoviti.politicheagricole.it/catalogo.php>

in formato Excel) è da iscriversi nel patrimonio informativo e quindi deve risultare come entità del Catalogo delle Banche Dati.

Affinché i dati possano essere considerati “materia prima” pregiata, la qualità più importante che essi devono possedere è rappresentata dal valore che un determinato settore o soggetto economico percepisce come significativo per il proprio *business* o per la propria attività istituzionale. In questo senso, i dati più pregiati sono quelli ad alto contenuto strategico, come i dati prodotti nell’ambito della Ricerca Scientifica.

Per la stretta relazione esistente tra i settori Agricolo ed Agro-alimentare e la qualità di vita e benessere dell’uomo e del pianeta, i dati relativi sono evidentemente di grandissimo valore. Sotto questa prospettiva, il settore della Ricerca e sperimentazione in Agricoltura, con particolare riferimento a quella pubblica, può essere un *driver* fondamentale per la *Data Economy*.

Valorizzazione del patrimonio informativo del CRA in modalità Open Data

In questo solco si inquadra l’impegno del CRA ad adottare politiche e strumenti finalizzati a facilitare l’accesso, la diffusione e il riuso delle informazioni e dei dati prodotti nell’ambito delle proprie attività di Ricerca e sperimentazione in Agricoltura, tenendo comunque conto degli aspetti legati alla tutela della riservatezza e alla salvaguardia della proprietà intellettuale.

Anche all’interno del CRA, gran parte dell’attività svolta viene ormai supportata da computer che gestiscono una notevole quantità di dati che sono parte, e in molti casi i risultati stessi, della Ricerca e sperimentazione in Agricoltura. Un primo censimento, tuttora in corso, ha già censito oltre 50 banche dati di estremo valore che già il CRA rende disponibili attraverso siti *web* dedicati. La pubblicazione degli stessi dati secondo i paradigmi Open Data ne aumenterà enormemente la fruibilità e la visibilità.

Attraverso il processo produzione, pubblicazione e diffusione delle informazioni e dei dati in modalità *Open Data*, il CRA ravvisa quindi la concreta possibilità di valorizzare il proprio patrimonio informativo, ovvero di dare il debito risalto alle innumerevoli attività di Ricerca e sperimentazione che troppo spesso rimangono patrimonio di una ristretta cerchia di operatori specializzati.

Il CRA decide quindi di aderire all’iniziativa di diffusione del proprio patrimonio informativo secondo i paradigmi Open Data al fine di agevolarne la diffusione e la valorizzazione, in linea con le raccomandazioni nazionali, comunitarie ed internazionali.

Agenda

L’Ente dovrà ora trasformare le dichiarazioni d’intenti contenute nel presente documento in fatti concreti seguendo un percorso ragionevolmente perseguibile e sostenibile.

Di seguito vengono elencate le prime azioni che il CRA dovrà avviare per dare attuazione alle linee strategiche sopra delineate:

- 1. Creazione dell'area "Valorizzazione del patrimonio informativo" all'interno della sezione "Amministrazione trasparente" del Portale Web del CRA.**
- 2. Realizzazione delle "Linee guida per la produzione di *Open Data* per l'Agricoltura".**
- 3. Creazione e pubblicazione dell'"Ontologia del CRA" (definizione di termini e relazioni in modalità comprensibile dalle macchine).**
- 4. Realizzazione del Portale *Open Data* per pubblicare sul *Web* i *dataset Open Data* del CRA.**
- 5. Supporto alle Strutture di ricerca per agevolare la pubblicazione dei dataset disponibili in modalità *Open Data*.**
- 6. Supporto alle iniziative che il MIPAAF intraprenderà in attuazione degli impegni assunti dall'Italia in sede G8.**
- 7. Collaborazione alle altre iniziative che il MIPAAF intenderà intraprendere.**

Ulteriori iniziative, in linea con questa strategia, saranno successivamente adottate secondo le necessità emergenti, anche al fine di cogliere le opportunità di sostegno finanziario che sorgeranno a livello nazionale e comunitario.