



IL PROGETTO “MAGAZZINI DEL DIGITALE”: PROVE DI SPERIMENTAZIONE

Giovanni Bergamin
Biblioteca Nazionale Centrale, Firenze

Abstract

Il termine *magazzini del digitale* è stato usato per la prima volta in occasione del progetto NEDLIB[1], poiché "il sistema di deposito delle pubblicazioni elettroniche è pienamente paragonabile a quello "reale" delle nostre biblioteche nazionali".

Il progetto "magazzini del digitali" – promosso dalla Fondazione Rinascimento Digitale e dalla Biblioteca Nazionale Centrale di Firenze – ha per obiettivo la realizzazione di un esperimento estensivo di archiviazione, con una ingente quantità di dati digitali.

L'architettura del progetto è basata sull'esperienza di due grandi archivi digitali già esistenti: Google e Internet Archive[2].

Considerando il fatto che "i difetti dei componenti sono la norma piuttosto che l'eccezione" la tolleranza ai difetti dell'intero sistema è fondata su:

- componenti hardware semplici e correnti, non legati a un particolare fornitore, che possano essere sostituiti facilmente (in pratica, personal computer con dischi SATA);
- replicazione dei dati attuata con strumenti open source di sincronizzazione dei dischi;
- software open source (sistema operativo e programmi gestionali);
- architettura distribuita comprendente due siti gemelli (Biblioteca Nazionale di Firenze e Biblioteca Nazionale di Roma) e un archivio di sicurezza (dark-archive).

[1]Johan Steenbakkers. *The NEDLIB guidelines*, 2000. <http://nedlib.kb.nl/results/NEDLIBguidelines.pdf>

[2] Internet Archive. Large Scale Data Repository: Petabox <http://www.archive.org/web/petabox.php>; Sanjay Ghemawat, Howard Gobioff e Shun-Tak Leung. The Google File System 2003. <http://labs.google.com/papers/gfs-sosp2003.pdf>