I SISTEMI DOCUMENTARI IN AMBIENTE DIGITALE

MARIA GUERCIO

Università di Urbino

m.guercio@mclink.it

II MODULO

I TEMI

- Gestione informatica dei documenti e gestione dei contenuti: convergenza e sovrapposizione
- Le criticità della transizione: la conservazione dei documenti informatici:
 - requisiti di autenticità
 - metodi e strumenti per la tenuta a lungo termine
- Le criticità della transizione: strumenti e iniziative di supporto

GESTIONE INFORMATICA DEI DOCUMENTI E GESTIONE DEI CONTENUTI: CONVERGENZA E SOVRAPPOSIZIONE

EDMS/ERMS: IL SISTEMA DOCUMENTARIO ARCHIVISTICO IN AMBIENTE DIGITALE

- Dalle linee guida europee (studio Moreq: www.ispo.cec.be/ida):
 - Nella creazione di sistemi documentari informatici, è necessario distinguere tra informazione, documento (document) e documento archivistico (record)
 - È necessario sviluppare sistemi informatici che:
 - formino (acquisiscano, gestiscano, utilizzino) e conservino documenti e risorse informative con le modalità, gli strumenti e le procedure specifiche di ciascuna tipologia di fonte, con particolare attenzione ai documenti archivistici che richiedono attenzione ai fini dell'autenticità e dell'affidabilità della fonte
 - integrino in un insieme organico e strutturato tutti gli elementi informativi disponibili sia all'interno che all'esterno del sistema documentario considerati utili al processo decisionale.

GESTIONE DEI CONTENUTI

- I sistemi di content management sono un insieme di attività e di prodotti software orientati ad organizzare i contenuti di un sistema informativo (non solo documentario) al fine di assicurare condizioni di interoperabilità e riuso dei contenuti tra sistemi o parti eterogenee di un sistema
- Hanno obiettivi vicini a quelli tradizionali (o rinnovati) dell'ERMS

GLI OBIETTIVI DEI CMS

- sviluppo di condizioni che garantiscano il riuso e l'integrazione di informazioni di provenienza diversa
- recupero efficiente dell'informazione
- tenuta del sistema informativo, dei suoi contenuti e della sua struttura
- facilitazione delle attività necessarie alla produzione, organizzazione e riuso dei contenuti mediante il trattamento specializzato dei metadati

GESTIONE DEI DOCUMENTI E CONTENUTI. CONSIDERAZIONI PRELIMINARI

- La sovrapposizione tra ERMS e CMS è spesso inevitabile ma non necessariamente negativa, soprattutto se dà luogo a processi guidati e consapevoli di reciproco riconoscimento, integrazione e convergenza. A tal fine sono necessarie alcune considerazioni:
 - Sistema informativo e sistema documentario costituiscono risorse integrate sempre meno distinguibili in fase di formazione eppure sempre più distinte nel successivo processo di tenuta e conservazione
 - La crescente interazione tra produzione e comunicazione del documento (pubblicazione sul web, sviluppo di intranet aziendali, lavoro cooperativo) comporta il rischio di un processo di indifferenziazione, appiattimento e confusione dei prodotti documentari e informativi e di perdita di qualità per la rapidità/immediatezza dei bisogni comunicativi
 - L'utilizzo di XML come meta-linguaggio di comunicazione e condivisione è un requisito significativo (sebbene richieda notevoli investimenti per la riconversione di contenuti legacy e per l'assenza di strumenti di mercato in grado di creare contenuti XML

GESTIONE DEI DOCUMENTI E CONTENUTI. LE POTENZIALITÀ DEI CMS

- predisposizione di schemi predefiniti di tipologie documentarie disponibili a utenti non esperti
- definizione e documentazione di politiche e di parametri di sicurezza informatica fondata sulla individuazione di ruoli utente e sulla identificazione e gestione di responsabilità specificate
- implementazione di sistemi di WFMS
- integrazione con i database e i sistemi informativi esistenti
- utilizzo di sistemi di autenticazione degli utenti
- gestione dei metadati
- trattamento flessibile delle informazioni a fini di riuso per finalità diversificate ("write once, publish many times"

GESTIONE DEI DOCUMENTI E CONTENUTI. LE RAGIONI DI UN RITARDO CULTURALE

- L'insufficienza di qualità dei contenuti disponibili nei sistemi documentari informatici a fronte di utenti sempre più esigenti ma anche distratti costituisce una delle ragioni per cui il mercato dei sistemi di EDMS e ERMS manifesta un considerevole ritardo rispetto ad altri settori dell'ICR
- Il perdurare di un gap culturale è legato:
 - all'incomprensione delle potenzialità che le tecnologie offrono per lo sviluppo di sistemi informativi documentari avanzati
 - all'offerta insufficiente di soluzioni organizzative e tecnologiche integrate da parte delle diverse componenti tecniche (archivisti e informatici)

GESTIONE DEI DOCUMENTI E CONTENUTI NON ESISTONO SCORCIATOIE

■ L'attesa di tecnologie mature che rendano superflua la fatica quotidiana dell'ordinamento razionale e sistematico delle informazioni documentarie prodotte e acquisite dalle organizzazioni sarà a lungo illusoria come dimostrano i modesti risultati ottenuti da progetti nordamericani o europei orientati allo sviluppo di sistemi informativi auto-organizzati, auto-aggiornati, auto-esplicativi.

GESTIONE DEI DOCUMENTI E CONTENUTI. L'INTEGRAZIONE NECESSARIA

- La spinta alla convergenza delle risorse documentarie determina la necessità di un processo duplice:
 - la specializzazione e il miglioramento degli strumenti di ERMS
 - lo sviluppo di sistemi e componenti software orientati all'integrazione dei contenuti (e-portal, groupware, workflow, data warehousing, search engine, messaging email, ecc.)
 - il riconoscimento che la classificazione d'archivio offre uno strumento avanzato più vicino alla gestione dei contenuti che alla semplice gestione dell'informazione, data la complessa analisi sintattica, semantica e funzionale che richiede
 - l'individuazione attenta delle criticità dei sistemi documentari informatici in questa lunga fase di transizione

CRITICITÀ DELLA FASE DI TRANSIZIONE

- sistemi ibridi
- conservazione a lungo termine di documenti elettronici a costi accettabili

LA CONSERVAZIONE DEI DOCUMENTI INFORMATICI

PERCHÉ LA CONSERVAZIONE È UN PROBLEMA CRUCIALE

- E' il nodo cruciale dell'automazione in campo documentario perché richiede un cambiamento significativo rispetto alle attività tradizionali: la conservazione in ambiente digitale è una funzione attiva e continua nel tempo per la quale non c'è ancora sufficiente esperienza, né un'adeguata analisi concettuale.
- I tempi degli interventi per il mantenimento della memoria si sono accorciati.
- La diversificazione dei prodotti non consente soluzioni univoche.
- La fragilità dei supporti e la facilità nella manipolazione richiedono investimenti significativi in termini di controllo dei depositi e sicurezza

... IRRISOLTO

- L'obsolescenza è un fenomeno irreversibile e ambivalente: per affrontarla non si sono ancora individuati metodi condivisi, univoci, regolamentati
- L'indifferenza del mercato (e qualche volta anche del legislatore e degli amministratori) per le esigenze conservative di lungo periodo è ampiamente testimoniata dall'assenza di standard internazionali, dall'insufficienza delle norme nazionali e delle procedure interne alle amministrazioni e, in Italia, dalle ridottissime risorse per la ricerca.

...E COMPLESSO

E' insufficiente garantire l'integrità del contenuto

- E' indispensabile la tenuta:
 - delle caratteristiche strutturali,
 - delle informazioni di provenienza e dei metadati di descrizione,
 - delle modalità di visualizzazione
- E' opportuno assicurare
 - il mantenimento della forma originaria, ma anche
 - l'accesso dinamico e interattivo alle fonti

DA CHE PARTE COMINCIARE

- dalla consapevolezza che sono necessari:
 - metodi e contenuti scientifici rigorosi nei diversi settori di appartenenza
 - programmi di formazione permanente di operatori preparati a una dimensione che muta in continuazione proponendo sfide sempre più impegnative,
 - determinazione nel superare difficoltà organizzative e nel promuovere soluzioni innovative basate sul coordinamento e sull'integrazione

La cooperazione è una risorsa, i cui costi (da non sottovalutare) hanno comunque risvolti positivi in termini di formazione permanente e qualificazione delle risorse professionali interne.

SUL PIANO TEORICO:

- dalla natura e dalla funzione dei documenti, dalla identificazione degli elementi che ne garantiscono l'autenticità, cioè identità e integrità (progetto InterPARES),
- dai metodi e dalle architetture sviluppate per assicurare, gestire e recuperare (nonostante l'obsolescenza e la frammentazione delle soluzioni applicative) contenuti, strutture formali, relazioni documentarie e di contesto che assicurano alle fonti documentarie significato e valore di testimonianza senza rinunciare all'efficiente ed efficace gestione e uso delle risorse digitali nell'attività corrente di chi le produce e le utilizza (progetto SDSC/NARA/NPACI)

SUL PIANO ORGANIZZATIVO:

- da iniziative "politiche" di sensibilizzazione indirizzate al legislatore, ai produttori, all'opinione pubblica, che si traducano in raccomandazioni, in linee guida, normativa
- dal riconoscimento da parte degli istituti di produzione e conservazione delle risorse digitali della centralità dei problemi organizzativi:
 - adozione di procedure adeguate e di standard
 - definizione di regole condivise
 - individuazione di responsabilità certe e riconosciute
- dalla valutazione, analisi e contenimento dei costi in relazione agli obiettivi e ai metodi possibili o necessari
- dalla riqualificazione "di massa" del personale tecnico mediante programmi di formazione permanente a distanza che richiedono aggiornamento dei contenuti e della didattica

LE RESPONSABILITA' SONO CONDIVISE E PRECOCI

- E' necessario (per contenere i costi e garantire i risultati) intervenire precocemente sin dalla formazione dei documenti
 - Larga parte delle informazioni che garantiscono l'accesso e la verifica dell'autenticità sono disponibili solo nella fase attiva di produzione e gestione delle risorse digitali (es. i sistemi di classificazione, i metadati di contesto, lo schema logico dei db, la documentazione delle applicazioni, ecc.).
- Sono a grave rischio tutti quei documenti che non si producono in un contesto organizzativo definito e, quindi, controllato
 - la documentazione di natura scientifica e personale,
 - I sistemi elettronici dinamici e interattivi (sistemi web)

LE RESPONSABILITA' DEI PRODUTTORI

- E' indispensabile riconoscere (in sede normativa? Nelle politiche nazionali per la conservazione?) la responsabilità dei produttori, che tuttavia vanno sostenuti con interventi diversificati:
 - Standard e linee guida
 - Training e best practice
 - Interventi di negoziazione/sensibilizzazione oltre che normativi (il deposito legale copre non più del 30% della produzione digitale destinata alla conservazione permanente, i documenti informatici protocollati e classificati costituiscono una percentuale ridotta dell'archivio anche nel settore pubblico)

LA RICERCA RICHIEDE INFRASTRUTTURE ADEGUATE

In particolare:

- risorse dedicate,
- laboratori permanenti e centri di competenza di livello internazionale,
- condivisione delle esperienze di ricerca e delle soluzioni applicative,
- continuità degli investimenti
- definizione di strategie di ricerca capaci di adattarsi all'evoluzione tecnologica

I RISCHI DA EVITARE

- la duplicazione delle iniziative
- la dispersione dei risultati per l'incapacità/l'impossibilità di convogliare soluzioni e materiali in canali di comunicazione efficaci

I FILONI DA SVILUPPARE

- Discussione e approfondimento delle basi teoriche e metodologiche per la conservazione permanente (ad esempio, il progetto InterPARES, < www.interpares.org >)
- Ricerca e sperimentazione (in ambiente cooperativo) di soluzioni avanzate (ad esempio, il progetto NPACI e SDSC: < www.npaci.edu.DICE/Pubs > o la proposta di laboratorio virtuale degli Archivi nazionali di Washington)
- Coordinamento operativo tra le istituzioni preposte alla conservazione del patrimonio culturale (il progetto MINERVA)
- Analisi, valutazione e comunicazione delle fonti informative e dei risultati utili ai responsabili per la conservazione (ad esempio, il progetto ERPANET <www.erpanet.org>)

LA VERIFICA DELL'AUTENTICITA'

- La conservazione a lungo termine in ambiente elettronico può assicurare esclusivamente la tenuta di copie autentiche di componenti digitali in grado di riprodurre (a richiesta dell'utente) copie autentiche di documenti informatici
- Il problema dell'autenticità è centrale
 - per i documenti attivi oggetto di migrazione tecnologica
 - per il mantenimento dei documenti versati negli istituti di conservazione: le condizioni di autenticità devono essere verificabili
 - per i documenti nella fase critica di trasferimento dall'ambiente di produzione originario a quello di consultazione a fini di ricerca
- La verifica dell'autenticità di un documento è possibile solo ricostruendo la storia del documento medesimo a condizione, quindi, che il documento ne abbia mantenuto le tracce (come avviene nel caso di documenti cartacei durevoli e stabili): è indispensabile perciò mantenere la **documentazione** relativa agli interventi di migrazione effettuati nel tempo e ai trattamenti subiti.

LA GARANZIA DELL'ACCESSIBILITÀ

- Mantenere, a costi accettabili, la possibilità di accesso e la fruizione efficiente implica la definizione di metodi per affrontare l'evoluzione delle tecnologie
- Non ci sono metodi oggi accettati che non implichino modifiche al flusso di bit dei documenti

IL RUOLO DI XML PER FORMARE PERSISTENT OBJECT

- Gli obiettivi operativi del progetto SDSC/NARA/NPACI:
 - predisporre un prototipo di strumento software indipendente dalle piattaforme, sufficientemente robusto e flessibile, basato sull'utilizzo di XML in quanto standard emergente (e promettente) per la rappresentazione e lo scambio informatico sul web
 - garantire la scalabilità dei prodotti con specifica attenzione alle esigenze di trusted digital repository di medie e piccole dimensioni.

UNA TEORIA DEI METADATI PER LA CONSERVAZIONE DIGITALE

- Identificare gli attributi/metadati relativi al documento e alle aggregazioni archivistiche e definire i pacchetti per il loro trattamento
- Descrivere i metadati per ciascun oggetto digitale (tipologia, formati, protocolli di acquisizione, metadati specifici di dominio, informazioni di presentazione), a livello di sistema (caratteristiche del sistema di memorizzazione, controlli di accesso e di audit, localizzazione, autenticazioni e cifrature)

PROGRAMMI E STRUMENTI PER SOSTENERE LE CRITICITA'

La complessità dei problemi determinati dall'obsolescenza richiede un impegnativo e mirato sostegno di lungo periodo (**programmi coordinati vs progetti**):

- per rafforzare (=rendere possibile) la cooperazione internazionale e la partecipazione dei ricercatori europei a indagini promettenti
- per sostenere la diffusione di informazioni qualificate e favorire l'aggiornamento professionale in un settore in cui è difficile orientarsi tra la grande quantità di fonti e materiali
- per avviare progetti mirati che:
 - coinvolgano ambienti disciplinari diversi finora poco attenti alle esigenze della conservazione permanente,
 - garantiscano una effettiva attività di sperimentazione,
 - includano la partecipazione di soggetti privati (società di informatica, settori industriali avanzati) realmente disposti ad impegnarsi in attività di ricerca
 - predispongano schemi e materiali per lo sviluppo di linee guida e raccomandazioni
- per favorire l'indispensabile sostegno "politico" a interventi complessi la cui necessità è scarsamente percepita dai governi: la conservazione è un processo e presuppone responsabilità molteplici

ERPA*NET*: UN'INFRASTRUTTURA PER LA COMUNICAZIONE

- Electronic
- Resource
- Preservation and
- Access
- NETwork



LE FINALITÀ DI ERPA*NET*

- superare l'isolamento (soprattutto delle istituzioni di piccole dimensioni) e garantire la multidisciplinarietà delle iniziative
- coinvolgere come nodi di una rete europea i centri di ricerca e le istituzioni che conservano il patrimonio culturale digitale, ma anche le società di informatica, il mondo economico, la pubblica amministrazione al fine di scambiare e condividere conoscenze e idee rendendole disponibili anche a coloro che non partecipano attivamente ai progetti di ricerca
- promuovere lo sviluppo e l'uso di standard e best practice

I CONTENUTI CONCRETI DI ERPANET

- Rendere disponibili documenti di ricerca, linee guida, standard significativi (cernita e valutazione delle fonti informative) mediante schede descrittive anche di sintesi
- Individuare i nuovi sviluppi del settore informatico utili ai fini della conservazione digitale
- Sviluppare e mettere a disposizione 60 studi di casi di best practice
- Sviluppare schemi di metadati e insiemi di requisiti funzionali
- Sviluppare strumenti e linee guida per almeno quattro aspetti principali: gestione del rischio, analisi dei costi, selezione di tecnologie e metodi, definizione di politiche generali e istituzionali
- Organizzare 6 workshop per esperti del settore e 6 seminari di aggiornamento (cfr tabella sul sito): Toledo (22-23 giugno 2002) sulla digitalizzazione), Urbino (9-11 ottobre 2002) sui metodi per la conservazione di documenti dinamici e interattivi, Roma (autunno 2003) sulle politiche per la conservazione digitale.

QUALCHE CONCLUSIONE

- La conservazione a lungo termine in ambiente elettronico può assicurare esclusivamente la tenuta di copie autentiche di componenti digitali in grado di riprodurre (a richiesta dell'utente) copie autentiche di documenti informatici
- IL problema dell'autenticità è centrale
 - per i documenti correnti oggetto di migrazione tecnologica
 - per il mantenimento e la possibile verifica dell'autenticità dei documenti versati negli archivi storici
 - per i documenti nella fase di trasferimento dall'ambiente di produzione originario a quello di consultazione a fini di ricerca
- La verifica dell'autenticità di un documento è possibile solo ricostruendo la storia del documento medesimo a condizione, quindi, che il documento abbia mantenuto le tracce della sua storia (come avviene nel caso di documenti cartacei durevoli e stabili): è indispensabile perciò mantenere la documentazione relativa agli interventi di migrazione effettuati nel tempo.

LA CONSERVAZIONE DEI DOCUMENTI INFORMATICI. METODI E SOLUZIONI

- La cooperazione internazionale è un requisito per sviluppare linee guida, raccomandazioni e soluzioni tecniche idonee
- Le ricerche e le sperimentazioni in corso non hanno ancora dato risultati univoci e operativi con alcune eccezioni (OAIS, InterPARES, progetto NPACI)
- La normativa tecnica sulla conservazione dei documenti informatici delle pp.aa. (ai sensi dell'art. 18 comma 3 del dpr 513/97) non ha ancora visto la luce
- Nella normativa italiana c'è molta confusione tra archiviazione sostitutiva e conservazione

INSUFFICIENZA DELLE NORME PROPOSTE

- Le norme hanno un andamento non lineare. In alcuni casi si caratterizzano per la genericità delle definizioni e delle disposizioni, in altri il dettaglio tecnologico finisce per complicare e rendere inattuabile e incomprensibile l'applicazione (ad esempio nel caso delle regole sull'archiviazione ottica)
- Manca una struttura concettuale chiara, condivisa e coerente, della cui rilevanza non c'è consapevolezza
- Il legislatore ha avuto fretta di completare il quadro normativo e ha erroneamente ritenuto di poter affrontare la conservazione di lungo periodo come un "semplice" problema di migrazione e di ridurre il problema della verifica dell'autenticità nel tempo a un problema di autenticazione nella fase di formazione, trasmissione o riproduzione.

LE CRITICITÀ DELLA TRANSIZIONE: INTERVENTI E STRUMENTI DI SUPPORTO

I RITARDI DELLE AMMINISTRAZIONI E DELLE IMPRESE

- I soggetti produttori non sono preparati ad affrontare le nuove sfide:
 - tendono ad affidarsi a risorse esterne,
 - sperano in soluzioni informatiche possibilmente sperimentate altrove,
 - non dispongono di profili professionali qualificati in questi settori
- I dirigenti sottovalutano la complessità di questa fase di informatizzazione soprattutto in ambito documentario

LE STRATEGIE POSSIBILI DI UNA FASE DI TRANSIZIONE

- Le soluzioni vanno trovate in modo graduale e nella sperimentazione di programmi che prevedano:
 - l'archiviazione sostitutiva nel rispetto della normativa vigente e di programmi di migrazione in relazione ai piani di classificazione e conservazione (soluzione organizzativa, più che tecnologica)
 - l'adozione di procedure che determinino la produzione di copie di sicurezza o la conservazione a lungo termine dei documenti su supporti di stabilità comprovata (es. microfilm)
- E' comunque opportuno produrre documenti informatici nel caso di materiali documentari di breve durata
- La scelta di intervenire su larga scala richiede l'adozione di sistemi di classificazione dettagliati che includono procedure per la conservazione e la selezione dei documenti

LE TRASFORMAZIONI ORGANIZZATIVE DEL SISTEMA DI PROTOCOLLO INFORMATICO

- Nuove funzioni di definizione, coordinamento e controllo con obiettivi di:
 - facilitazione del decentramento delle funzioni operative
 - uniformità dei criteri e delle procedure
 - disponibilità delle informazioni documentali
 - miglioramento delle funzionalità di inserimento dati e di reperimento
 - diminuzione delle fasi ripetitive e ridondanti
 - garanzia per la sicurezza e integrità del sistema documentale
- Il sistema tende ad articolarsi in strutture decentrate coordinate (archivi gestionali) che "concorrono alle attività di registrazione di protocollo e archiviazione".
- Sono necessarie linee guida e direttive settoriali, assistenza e sostegno

CENTRALITA' DELLA FORMAZIONE

Nodo centrale è l'esistenza di personale in grado di:

- formare il sistema documentario in ambiente digitale
- dialogare con gli specialisti di altre discipline
- seguire l'evoluzione tecnologica

Nel caso delle istituzioni archivistiche si richiede inoltre:

- una strategia consapevole
- la condivisione degli obiettivi da parte di tutta la struttura

LE INIZIATIVE DI ORIENTAMENTO IN PROGRAMMA

- diffusione di best practice (www.aipa.it: gestione di un sito web articolato e aggiornato)
- elaborazione di strumenti di orientamento (Gedoc2)
- creazione di un gruppo di lavoro presso la Scuola superiore della p.a. per:
 - definire <u>linee guida per la</u> gestione informatica dei documenti e uno schema condivisibile per le funzioni di auto-amministrazione,
 - elaborare uno schema di manuale di gestione
 - predisporre <u>materiali riutilizzabili</u> ai fini della formazione, a partire da uno <u>metodologia per la classificazione d'archivio</u>
- Sviluppare piani di formazione coerenti (progetto AIPA, progetti SSPA, nuovi corsi universitari orientati alla formazione di archivisti contemporanei)

FINALITA' DEI PROGETTI DELLA SCUOLA SUPERIORE DELLA P.A. PER LA GESTIONE ELETTRONICA DEI DOCUMENTI

- Sostenere le amministrazioni pubbliche nello sforzo di riqualificare le risorse umane interne destinate a un ruolo di direzione (responsabili e funzionari dei Servizi per la gestione informatica dei documenti)
- Promuovere la cooperazione nel campo della ricerca di settore e consentire un confronto di studio e di sperimentazione tra soggetti produttori di documenti e interlocutori istituzionali (amministrazioni nazionali e amministrazione archivistica, Autorità per l'informatica, Dipartimento per la funzione pubblica)

I CORSI DI FORMAZIONE SSPA (2000-2002)

- Corsi di base sulla gestione elettronica degli archivi (2000-2002): corsi erogati per 300 dipendenti
- Corsi di base sulla gestione elettronica degli archivi orientati a utenze specialistiche (moduli personalizzati per settori amministrativi)
- Corsi di specializzazione sulla classificazione d'archivio (2002): corsi-laboratorio per la progettazione di sistemi di classificazione in ambiente digitale
- Corsi di alta formazione per formatori (2002)

QUALCHE INDIRIZZO

- http://.erpanet.org
- http://www.interpares.org (progetto InterPARES)
- http://www.ispo.cec.be/ida (progetto Moreq)
- http://jitc.fhu.disa.mil/recmgt/ (Standard 5015.2 "Design Criteria Standard For Electronic Records Management Software Applications" US Department of Defense)
- http://ccsds.org/RP9905/RP9905.html (proposta standard OAIS)
- http:// sdsc.edu/NARA e http://www.npaci.edu.DICE/Pubs
 (progetto del National Archives di Washington e del Supercomputer Center di S.Diego)
- <u>http://www.sspa.it</u> (Scuola superiore per la pubblica amministrazione: materiali didattici, pagine web sulla ricerca) e <u>http://www.aipa.it</u> (normativa, pagine web sul protocollo)