

# Il modello di riferimento per un Sistema informativo aperto per l'archiviazione

Brian F. Lavoie<sup>1</sup>

I. Breve premessa.....	1
II. La genesi del modello OAIS .....	2
III. Il sistema informativo aperto per l'archiviazione (OAIS) .....	5
IV. L'ambiente OAIS .....	8
V. Il modello funzionale OAIS .....	12
VI. Il modello informativo OAIS .....	15

## I. Breve premessa

L'informazione digitale ha davvero comportato cambiamenti per tutti: per l'industria, i governi e il mondo accademico; per gli uomini d'affari, gli scienziati, gli ingegneri e gli studiosi di discipline umanistiche; per il singolo individuo, tanto sul posto di lavoro quanto a casa.

Una grande quantità di informazione digitale – testi, immagini, audio, video, pagine web, programmi informatici, database – viene prodotta, scambiata ed utilizzata in una varietà di contesti e per una miriade di

---

<sup>1</sup> Traduzione di *Open Archival Information System Reference Model: Introductory Guide / Lavoie, Brian F* (2004) reperibile in [http://www.dpconline.org/docs/lavoie\\_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf). Questa traduzione non vuole essere una traduzione letterale del testo: vengono qui tradotti con semplificazioni e adattamenti che facilitano la lettura i primi sei capitoli. Il modello OAIS è un modello generale: non si riferisce solo alle risorse digitali, ma a qualsiasi tipo di risorsa informativa si desidera conservare e rendere accessibile nel lungo periodo. Tuttavia - come sottolinea l'autore nella nota n. 5 - il modello OAIS si è diffuso come modello di riferimento per le risorse digitali e di solito viene studiato e applicato in questo contesto.

scopi. Le applicazioni della tecnologia digitale si avvalgono di una infrastruttura che comporta benefici condivisi e che si basa su:

- potenti capacità di ricerca e di recupero dell'informazione;
- distribuzione tramite la rete;
- illimitata possibilità di duplicazione;
- interoperabilità.

Gli effetti benefici dell'informazione digitale e le sfide che la accompagnano trascendono le persone, i sistemi e le comunità. Questo non è stato mai così evidente come nel caso della conservazione delle risorse digitali (dove *conservare* significa mantenere nel lungo periodo le possibilità di accesso e di fruizione all'informazione custodita).

Si sta consolidando sempre più la capacità sia di creare che di consumare l'informazione digitale, mentre si sta facendo strada in maniera più lenta la capacità *conservare* nel lungo periodo di questo tipo di informazione. La fragilità dei supporti e la rapida obsolescenza dei sistemi di archiviazione delle risorse digitali aggrava il problema: siamo costretti ad affrontare le problematiche relative alla *conservazione* in un orizzonte temporale relativamente breve.

Tutte le organizzazioni e comunità (culturali, governative, commerciali ecc) si stanno ponendo il problema dell'accessibilità nel lungo periodo delle risorse digitali in loro possesso anche se con gradi di consapevolezza diversi. La trasversalità delle problematiche della conservazione delle risorse digitali richiede un terreno comune di dialogo e di cooperazione interdisciplinare al fine di affrontare le sfide emergenti.

L'investimento nella ricerca in soluzioni per la conservazione delle risorse digitali in una comunità spesso produce un effetto onda che ha un impatto su molte altre comunità. Ed è proprio questo che è successo quando la comunità custode dei dati relativi alle esplorazioni dello spazio prese in conto il problema della conservazione dei suoi dati digitali.

## **II. La genesi del modello OAIS**

Il *Consultative Committee for Space Data Systems (Comitato consultivo per i sistemi dei dati spaziali)*, CCSDS, istituito nel 1982, è

un coordinamento tra le agenzie spaziali nazionali interessate allo sviluppo cooperativo di standard per la gestione dei dati relativi alla ricerca spaziale. Nel 1990, il CCSDS avviò un accordo di cooperazione con ISO, in base al quale le Raccomandazioni CCSDS - ad esempio, le soluzioni raccomandate e condivise dai suoi membri nel campo della gestione dei dati scientifici - sarebbero state sottoposte alle normali procedure ISO di revisione e di voto per diventare formalmente standard ISO.

Su richiesta dell'ISO, il CCSDS iniziò un lavoro con l'obiettivo di sviluppare standard per l'archiviazione a lungo termine dei dati digitali generati dalle missioni spaziali. Nell'indagine preliminare, il CCSDS non trovò un quadro di riferimento largamente accettato che potesse servire quale fondamento per l'attività di costruzione degli standard.

Mancava, ad esempio, un quadro comune di riferimento per:

- la descrizione delle funzioni basilari costituenti un sistema per l'archiviazione delle risorse digitali;
- la definizione degli attributi fondamentali degli oggetti informativi digitali da conservare.

In breve, non c'era una consapevolezza condivisa sui bisogni e sui requisiti per il mantenimento nel lungo periodo dell'informazione digitale. Un quadro di riferimento unificante che potesse colmare questa lacuna sarebbe estremamente utile per:

- favorire il dialogo e la collaborazione nella preparazione degli standard;
- identificare le aree che con più probabilità avrebbero beneficiato dallo sviluppo di questi standard.

In mancanza di un tale quadro di riferimento, il CCSDS stabilì che il suo primo passo sarebbe stato quello di crearne uno. Un seminario internazionale organizzato dal CCSDS nel 1995 confermò la validità di questa strategia con la proposta di sviluppare *"un modello di riferimento per un sistema informativo aperto per l'archiviazione"*. Il *modello di riferimento* avrebbe definito – con una terminologia non legata a un qualche specifico settore:

- le componenti funzionali di base di un sistema dedicato alla conservazione nel lungo periodo dell'informazione digitale;
- le interfacce interne ed esterne del sistema;
- gli oggetti gestiti;
- i requisiti minimi da soddisfare.

Una volta completo, il *modello* avrebbe rappresentato un quadro di riferimento completo e coerente per:

1. descrivere ed analizzare le problematiche relative alla conservazione delle risorse digitali;
2. fornire una solida base per l'attività successiva di definizione degli standard;
3. servire come un punto di riferimento per i fornitori interessati a costruire prodotti e servizi per la conservazione delle risorse digitali.

A partire dalle attività iniziali di sviluppo, il CCSDS fu subito consapevole che la rilevanza del *modello* andava ben al di là della comunità dei dati spaziali. Il *modello di riferimento* si sarebbe occupato delle questioni fondamentali relative alla conservazione nel lungo periodo delle risorse digitali - questioni che non sono legate a una specifica area. Per questo il CCSDS decise di rendere il processo di creazione del *modello* aperto a qualsiasi individuo o organizzazione interessati

Il *modello di riferimento* rappresenta:

- il terreno comune su cui consolidare la comprensione dei bisogni e dei requisiti relativi alla conservazione delle risorse digitali;
- l'opportunità di mettere insieme i frammenti di attività isolate di conservazione delle risorse digitali, unificandole in una individuazione condivisa (per quanto altamente concettuale) dei confini del problema.

Il *modello di riferimento* fu sviluppato attraverso un processo di bozze, riesame e revisione, aperto e iterativo; la partecipazione delle comunità interessate fu ottenuta con discussioni seminariali o con risposte scritte a richieste formali di commento. Versioni in bozza del *modello di riferimento* furono rilasciate nel maggio 1997 e nel maggio 1999; nel giugno 2000 il *modello* fu approvato e pubblicato come standard ISO in bozza. Dopo un periodo di riesame e revisione, il *modello di riferimento* fu approvato nel gennaio 2002 come standard internazionale ISO 14721.

### III. Il sistema informativo aperto per l'archiviazione (OAIS)

L'idea centrale del *modello di riferimento* è quella di un sistema informativo aperto per l'archiviazione (Open archival information system, OAIS). IL termine "aperto" si riferisce al fatto che come abbiamo visto il processo di creazione del *modello* è stato un processo - appunto - aperto a qualsiasi individuo o organizzazione interessati. Un sistema informativo per l'archiviazione è "un'organizzazione di soggetti e sistemi che ha accettato la responsabilità della conservazione dell'informazione e del mantenerla disponibile per una comunità di riferimento". Questa definizione dà particolare rilievo alle due funzioni primarie di un sistema per l'archiviazione:

1. conservare l'informazione - per esempio, assicurare la sua persistenza nel lungo periodo;
2. fornire l'accesso all'informazione archiviata, in un modo coerente ai bisogni degli utenti primari di un sistema OAIS, o "comunità di riferimento". Il concetto di "comunità di riferimento" sarà approfondito nella prossima sezione.

Il termine "archivio" viene associato a ben precise responsabilità; in particolare, un archivio conforme a OAIS deve:

- negoziare e accettare dai produttori le risorse informative da archiviare;
- ottenere il controllo dell'informazione archiviata a un livello tale da assicurare la conservazione nel lungo periodo;
- definire la "comunità di riferimento" dell'archivio;
- assicurare che l'informazione conservata sia comprensibile autonomamente dalla comunità degli utenti, e cioè che l'informazione possa essere usata dagli utenti senza ricorrere al produttore dell'informazione stessa.
- seguire linee di comportamento e procedure documentate in maniera tale da assicurare che le informazioni siano conservate con ragionevole sicurezza e possano essere distribuite come copie autentiche o rapportabili all'originale;
- rendere l'informazione conservata disponibile per la "comunità di riferimento"

La prima responsabilità di un archivio di tipo OAIS è di stabilire i criteri di selezione e di acquisizione. Una volta definito i "confini" della

raccolta d'archivio, occorre motivare adeguatamente i produttori/proprietari a depositare gli oggetti informativi individuati. Ma non si tratta semplicemente di acquisire la custodia di tali oggetti.

La seconda responsabilità enfatizza infatti la necessità per un archivio conforme al modello OAIS di ottenere il controllo dell'informazione archiviata a un livello tale da assicurare la conservazione nel lungo periodo. Per esempio, se un archivio OAIS deve creare una nuova versione dell'oggetto archiviato in modo tale che questo possa essere fruito con le tecnologie correnti, deve poterlo fare senza con questo violare i diritti di proprietà intellettuale

Un'altra responsabilità per un archivio di tipo OAIS è definire i confini della comunità degli utenti di riferimento. Il concetto di "comunità di riferimento" sarà esaminato nella prossima sezione ed è una precondizione per soddisfare un'altra delle responsabilità OAIS: assicurare che l'informazione sia conservata in una forma che sia autonomamente comprensibile dagli utenti. La produzione dell'informazione avviene sempre in un certo contesto, e spesso accade che la comprensione di tale contesto sia necessaria per la piena comprensione dell'informazione stessa. Per questo un archivio di tipo OAIS deve non solo conservare l'informazione, ma il contesto necessario ad assicurare che l'informazione sia "comprensibile", e in definitiva, "utilizzabile" nel tempo. Le informazioni relative al contesto comprendono

- una descrizione della struttura o del formato in cui l'informazione è archiviata,
- come e perché l'informazione è stata creata

Le due responsabilità finali dell'OAIS riguardano il processo di conservazione e la fruibilità dell'informazione archiviata. Un archivio conforme a OAIS deve:

- stabilire e documentare politiche e procedure chiare per svolgere le funzioni di conservazione dell'informazione in sua custodia.
- rendere disponibili i contenuti archiviati alla propria comunità di utenti di riferimento, attraverso l'implementazione di meccanismi di accesso e di servizi in grado di rispondere , per quanto possibile, ai bisogni e alle richieste degli utenti.

L'uso del termine OAIS, o ugualmente, del termine "archivio" in un contesto OAIS, implica:

- un sistema per l'archiviazione dedicato alla conservazione dell'informazione digitale (dove *conservare* significa mantenere nel lungo periodo le possibilità di accesso e di fruizione all'informazione custodita)
- l'assunzione con determinate modalità delle sei responsabilità obbligatorie prima elencate.

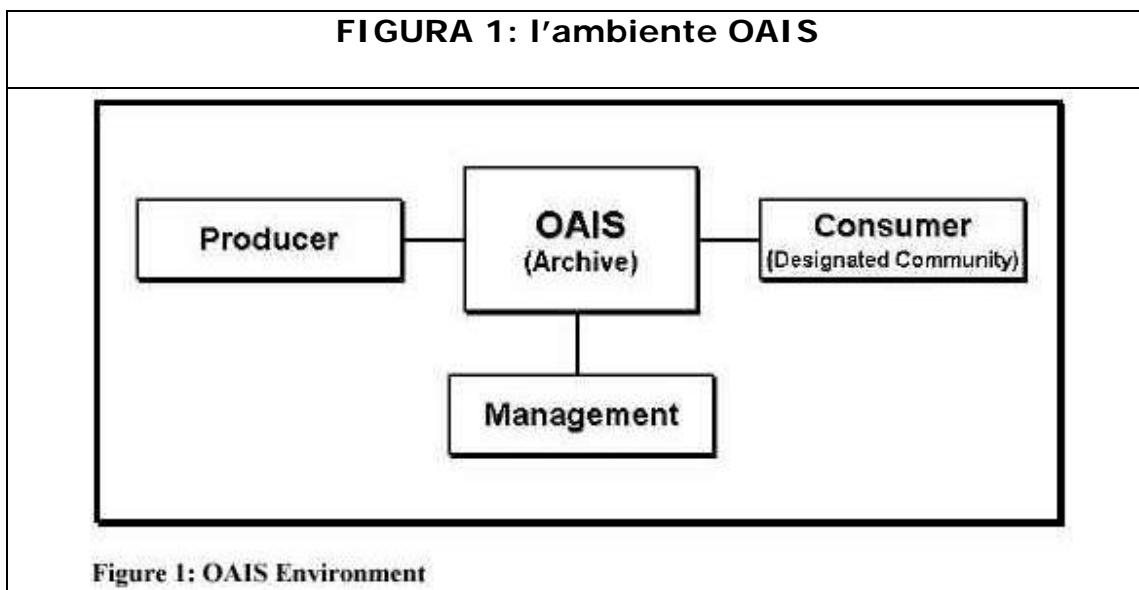
Il *modello di riferimento OAIS* si sviluppa in tre parti:

- la prima parte descrive l'ambiente esterno con il quale un archivio di tipo OAIS interagisce;
- la seconda parte descrive le componenti funzionali, o i meccanismi interni, che insieme rispondono alle responsabilità relative alla conservazione;
- la terza parte descrive gli oggetti informativi (information objects) che sono acquisiti, gestiti e messi a disposizione degli utenti.

Le prossime tre sezioni approfondiscono ognuna di queste parti.

#### IV. L'ambiente OAIS

Un archivio di tipo OAIS non opera in maniera isolata, ma si fa carico delle responsabilità relative alla conservazione ed all'accesso in un ambiente popolato da più soggetti chiave esterni. Un archivio di tipo OAIS deve operare in cooperazione con questi partecipanti esterni per adempiere alla propria missione. Il modello di riferimento identifica e descrive i soggetti esterni che costituiscono un ambiente OAIS, e definisce l'interfaccia tra questi soggetti e l'OAIS.



La figura 1 illustra l'ambiente OAIS. L'ambiente comprende tre distinti soggetti esterni ma interagenti con OAIS:

1. *Management (Gestione)*;
2. *Producer (Produttore)*;
3. *Consumer (Utente/consumatore)*.

Le responsabilità del *Management* comprendono la progettazione, la revisione e il rafforzamento, del quadro di riferimento della politica di alto livello che governa le attività OAIS.

Esempi delle funzioni espletate dal *Management* sono la pianificazione strategica, la definizione dei confini della raccolta archiviata dell'OAIS, la specificazione delle garanzie di conservazione che si offrono per gli oggetti affidati all'archivio. Il *Management* può essere anche la fonte di finanziamento per l'OAIS con il ruolo anche di supervisione, ovvero di riesaminare periodicamente le politiche e i risultati.

Occorre rilevare che il *Management* non è responsabile della gestione delle operazioni quotidiane richieste da un archivio di tipo OAIS.



Questa responsabilità è gestita da una componente funzionale interna dell'archivio (si veda la Sezione V).

Il secondo soggetto esterno che interagisce con l'OAIS è il *Producer* (il produttore o i produttori): individui, organizzazioni, o sistemi che depositano risorse digitali in un archivio OAIS per la conservazione nel lungo periodo. I produttori inviano a un archivio OAIS le risorse digitali da conservare e i relativi metadati attraverso un "processo di immissione" (*ingest process*), che accetta i dati inviati e li prepara per l'inclusione nel sistema per l'archiviazione. L'interazione tra l'OAIS e i produttori è spesso formalizzata e guidata da un "accordo per l'invio dei dati" (*submission agreement*), che stabilisce i dettagli dell'interazione quali il tipo di informazione da inviare, i metadati che il produttore si impegna a fornire e le modalità operative del trasferimento dall'archivio del produttore all'archivio OAIS.

L'OAIS interagisce anche con gli utenti (*consumer*), non solo con il Management e i produttori. Come suggerisce il nome (*consumer*), gli utenti sono individui, organizzazioni o sistemi che fruiscono delle informazioni conservate dall'OAIS. Il modello di riferimento precisa una classe speciale di utenti indicata come la "comunità di riferimento": il sottoinsieme degli utenti in grado di comprendere autonomamente l'informazione archiviata nella forma in cui è conservata e resa disponibile dall'OAIS. Infatti una delle responsabilità obbligatorie di un archivio OAIS è quella di conservare l'informazione in un modo che sia comprensibile agli utenti di riferimento.

Ad esempio se l'OAIS contiene studi o insiemi di dati relativi a una determinata disciplina, allora la "comunità di riferimento" può consistere di tutti gli individui che possiedono un certo livello di conoscenze in quell'area, che potrebbero usare quanto è stato archiviato per ottenere informazioni o come punto di partenza per la ricerca di base o applicata. Così, se il contenuto di un archivio OAIS consiste in bilanci, dichiarazioni dei redditi o in altri documenti finanziari di imprese commerciali, la "comunità di riferimento" potrebbe identificata con gli enti di controllo governativi e con i professionisti del mondo finanziario e contabile in grado di offrire sintesi e interpretazioni di queste informazioni.

In entrambi gli esempi, i contenuti dell'OAIS potrebbero essere liberamente accessibili da chiunque; in questo caso, gli utenti di un OAIS coinciderebbero con il pubblico in generale. Tuttavia solo chi possiede una conoscenza sufficientemente specializzata per usare le

informazioni archiviate senza l'assistenza di un esperto fa parte della "comunità di riferimento" dell'OAIS.

La "comunità di riferimento" non è definita ex post dalla natura dei contenuti dell'archivio; piuttosto, è la natura della "comunità di riferimento" che determina sia i contenuti dell'OAIS sia le modalità di conservazione.

Definire i confini della "comunità di riferimento" è un aspetto critico per un archivio conforme a OAIS. Come l'approfondimento nella Sezione VI renderà chiaro, più ampio è il confine della "comunità di riferimento", maggiori saranno i requisiti relativi ai metadati necessari per il mantenimento nel lungo periodo dei materiali digitali. La "comunità di riferimento" potrebbe estendersi quanto la popolazione; questo equivale a non richiedere nessuna particolare capacità o conoscenza specifica da parte degli utenti dell'informazione archiviata. Ma in questo caso, il compito di conservare l'informazione in una forma "autonomamente comprensibile" diventa ovviamente più difficile. Inoltre i confini della "comunità di riferimento" possono cambiare nel tempo: la "comunità di riferimento" può estendersi, come pure i suoi membri possono cambiare le proprie aspettative in merito all'accesso e all'uso dei contenuti di un archivio OAIS.

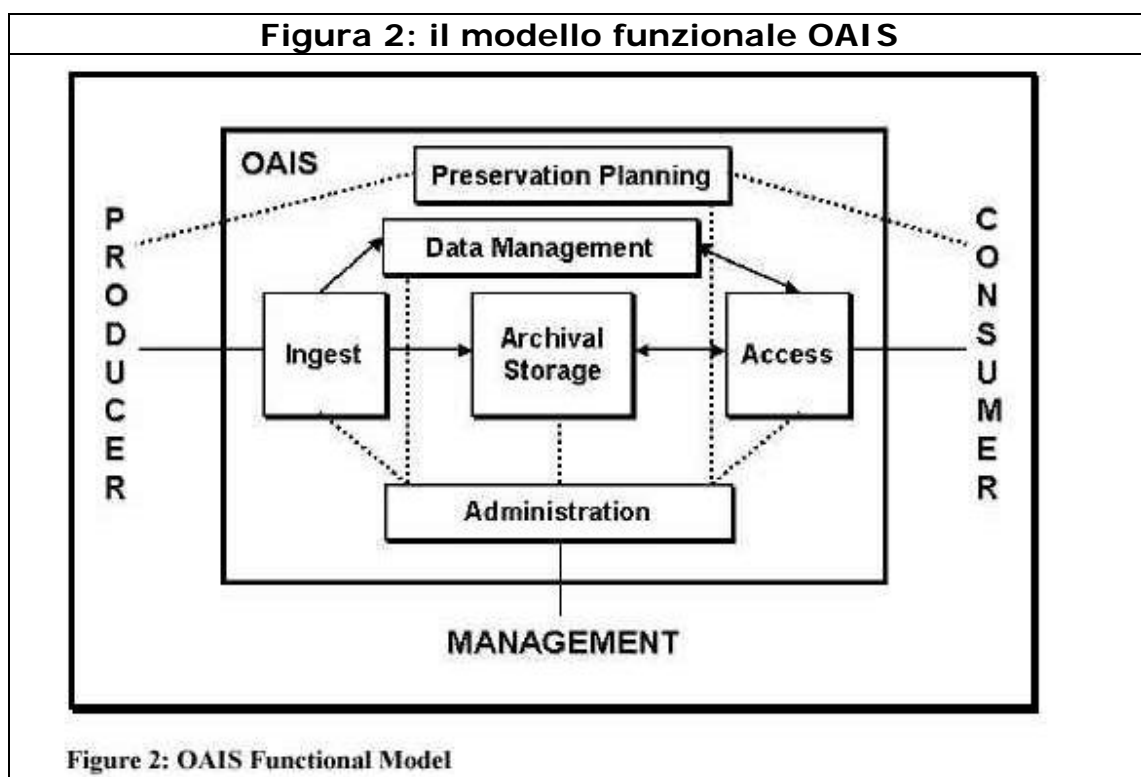
I concetti di *Management*, *Produttori*, *Utenti* e "comunità di riferimento", come quello di un *archivio OAIS*, rappresentano ruoli funzionali piuttosto che organizzativi. Tutti questi ruoli possono essere affidati a una singola struttura organizzativa, o distribuiti in diverse organizzazioni. Il punto chiave non è la separazione fisica di un ruolo da un altro, ma piuttosto, la separazione logica dei ruoli decisionali e degli interessi dei soggetti coinvolti.

E' utile un esempio per capire l'applicazione in pratica dell'ambiente OAIS. Il National Digital Archive of Datasets (NDAD) è una iniziativa di base nel Regno Unito, mirata a conservare i dati informatici prodotti dai dipartimenti e dalle agenzie del governo centrale britannico. In questo scenario, l'OAIS è il servizio di National Data Repository (NDR), un sistema di conservazione delle risorse digitali e di accesso a cura del Centro informatico dell'Università di Londra (University of London Computer Centre, ULCC). Il ruolo gestionale, il *Management*, comunque, risiede presso i National Archives del Regno Unito, che mantengono la custodia legale dei dati archiviati e svolgono un numero di funzioni di alto livello associate all'iniziativa NDAD, compreso il finanziamento di NDAD e la selezione di dati per la conservazione di lungo periodo. I produttori sono, naturalmente, i vari dipartimenti e le

agenzie del governo britannico che, come parte della loro missione organizzativa, producono dati informatici. I dati archiviati sono liberamente disponibili per chiunque disponga di un accesso web. Una visita al sito web di NDAD suggerisce che la "comunità di riferimento" si estende al pubblico generale: per utilizzare il database NDAD occorre semplicemente avere la capacità di usare un browser. Inoltre, i dati archiviati sono accompagnati da informazioni descrittive dettagliate che spiegano "chi, come e quando i dati sono stati creati". In breve i dati sono "autonomamente comprensibili" da un pubblico non specializzato (che conosca naturalmente la lingua inglese).

## V. Il modello funzionale OAIS

Il modello di riferimento identifica e descrive l'insieme fondamentale dei meccanismi con i quali un archivio di tipo OAIS adempie la sua missione primaria di conservare l'informazione a lungo termine e di renderla disponibile alla "comunità di riferimento". Questi meccanismi sono riassunti dal "modello funzionale" OAIS: una raccolta di sei servizi di alto livello, o componenti funzionali.



La seconda figura illustra il modello funzionale OAIS. La prima componente funzionale è l' "Immissione" (*Ingest*), l'insieme dei processi responsabili dell'accettazione delle risorse digitali inviate dai Produttori e della loro preparazione per l'inclusione nel sistema per l'archiviazione. Le funzioni specifiche svolte dalla componente funzionale "Immissione" comprendono:

- la presa in carico delle risorse digitali trasferite all'OAIS dal Produttore;
- la verifica che quanto ricevuto sia integro e completo;
- gli interventi che rendono gli oggetti ricevuti gestibili da OAIS

- l'estrazione e/o la creazione di metadati descrittivi per supportare la ricerca e le funzionalità di recupero;
- il trasferimento dell'informazione inviata e dei metadati associati al sistema per l'archiviazione.

In breve, la funzione di *"Immissione"* funge da interfaccia esterna dell'OAIS con i Produttori: si fa carico dell'intero processo di:

- accettazione della custodia dell'informazione sottoposta
- preparazione dei dati ricevuti per l'archiviazione.

La seconda componente funzionale dell'archivio di tipo OAIS è il *"Archiviazione dei dati" (Archival Storage)*. Questa è la parte del sistema d'archiviazione che gestisce l'immagazzinamento a lungo termine delle risorse digitali affidate all'OAIS. Più specificamente, la funzione di *Archival Storage* ha le seguenti responsabilità:

- affidabilità e funzionalità dei sistemi di immagazzinamento (storage);
- integrità e fruibilità a lungo termine delle sequenze di bit (bit stream) che compongono i dati conservati

Per adempiere a questa responsabilità, la componente *Archival Storage* intraprende periodicamente procedure quali l'aggiornamento dei supporti di archiviazione (refresh) e la migrazione dei formati. La funzione di *Archival Storage* implementa anche diversi meccanismi di salvaguardia:

- le procedure di verifica di errore
- le politiche di recupero da disastro (disaster recovery) per mitigare gli effetti di eventi catastrofici.

Infine, la componente *Archival Storage* recupera gli oggetti dai sistemi di archiviazione per rispondere alle richieste di accesso degli utenti. Si noti che la funzione dell' *Archiviazione dei dati* non ha una interfaccia esterna diretta; per interagire con l' *Archival Storage* occorre passare attraverso le altre componenti funzionali.

La *Gestione dei dati (Data Management)* è la terza componente funzionale di un OAIS. La funzione di *Data Management* mantiene i database dei metadati descrittivi che identificano e descrivono le informazioni archiviate; *Data Management* si occupa anche i dati amministrativi relative alle operazioni del sistema interno OAIS, come i dati relativi alle prestazioni del sistema e le statistiche di accesso. Le funzioni primarie del *Data Management* includono

- la manutenzione dei database di cui è responsabile;

- l'esecuzione di ricerche su questi database e la produzione di rapporti in risposta alle richieste provenienti da altre componenti funzionali dell'OAIS,
- l'aggiornamento dei database non appena arrivino nuove informazioni o quando l'informazione esistente venga modificata o cancellata.

Attraverso questi database la componente funzionale *Data Management* consente sia la ricerca ed il recupero del contenuto archiviato, sia lo svolgimento delle operazioni interne di un archivio OAIS.

La quarta componente funzionale di un OAIS è il *Preservation Planning* (*Pianificazione della conservazione*). Questo servizio è responsabile sia della progettazione della strategia di conservazione dell'OAIS sia della sua revisione in risposta a cambiamenti tecnologici riguardanti oggetti archiviati. Il servizio di *Preservation Planning* guarda all'ambiente esterno e cerca di intercettare tutti quei cambiamenti che potrebbero avere un impatto sulla capacità dell'OAIS di conservare e mantenere l'accesso alle informazioni in sua custodia (come le innovazioni nelle tecnologie di immagazzinamento e di accesso, o mutamenti nei confini o nelle aspettative della "comunità di riferimento"). In risposta a questi cambiamenti il servizio di *Preservation Planning* elabora le raccomandazioni per l'aggiornamento delle politiche e delle procedure dell'OAIS. La funzione di *Preservation Planning* rappresenta la salvaguardia dell'OAIS nei confronti di un ambiente in costante evoluzione in termini di utenti e tecnologie.

L'*Accesso* (*Access*) è la quinta componente funzionale di un archivio di tipo OAIS. Come suggerisce il nome, la funzione *Access* gestisce i processi ed i servizi attraverso i quali gli Utenti – e in special modo la "comunità di riferimento" – localizza, richiede e riceve gli oggetti che risiedono nel sistema per l'archiviazione dell'OAIS. I servizi tipici forniti da *Access* consistono in

- l'elaborazione delle richieste ricevute relative al possesso dall'OAIS: inoltro della richiesta al *Data Management* e la presentazione della risposta (per esempio, una serie di risultati) all'utente;
- il coordinamento tra il recupero dell'informazione la consegna del contenuto richiesto: l'inoltro della richiesta all'*Archival Storage*, preparazione dell'oggetto digitale per la consegna

*Access* è anche responsabile dell'implementazione di ogni meccanismo di controllo dell'accesso e della sicurezza associato con il contenuto archiviato. La funzione *Access* rappresenta l'interfaccia dell'OAIS con i

suoi Utenti (e la “comunità di riferimento”): pertanto è il meccanismo primario attraverso il quale l’OAIS ottempera alle proprie responsabilità.

La sesta e finale componente funzionale di un OAIS è l’Amministrazione (Administration). La funzione Administration è responsabile della gestione delle operazioni quotidiane dell’OAIS:

- coordinamento delle attività degli altri cinque servizi di alto livello dell’OAIS.
- l’interazione con i produttori (ad esempio la negoziazione degli accordi per la presentazione dei documenti),
- l’interazione con gli utenti (per esempio, assistenza),
- l’interazione con il Management (per esempio, l’implementazione e il mantenimento delle politiche e degli standard di archiviazione).
- supervisione dei sistemi di archiviazione e di accesso, del monitoraggio delle prestazioni di sistema, e del coordinamento degli aggiornamenti del sistema.

Administration è il punto centrale per le interazioni interne ed esterne dell’OAIS: comunica direttamente sia con gli altri cinque servizi di alto livello dell’OAIS – Ingest, Archival Storage, Data Management, e Access, sia con gli attori esterni – Produttori, Utenti e Management.

Riassumendo, l’OAIS comprende sei componenti funzionali (o servizi) di alto livello che insieme costituiscono i meccanismi con cui l’OAIS conserva le informazioni a lungo termine e le rende disponibili alla “comunità di riferimento”.

## **VI. Il modello informativo OAIS**

In aggiunta alla descrizione delle componenti funzionali di un archivio di tipo OAIS, il modello di riferimento fornisce anche una descrizione di alto livello degli *oggetti informativi (information object)* gestiti dall’archivio. Il modello informativo OAIS è costruito intorno al concetto di *pacchetto di informazioni (information package)*. Si tratta di una rappresentazione concettuale dell’informazione nelle sue fasi di vita: immissione, archiviazione e distribuzione. Un *pacchetto di informazione* consiste in un insieme logico composta da:

- l’oggetto digitale (ciò che deve essere conservato)
- metadati necessari a garantire conservazione ed accesso sul lungo periodo.

Ci sono tre importanti versioni del concetto di *information package*:

- per l'immissione (Submission Information Package o SIP);
- per l'archiviazione (Archival Information Package o AIP);
- per la distribuzione (Dissemination Information Package o DIP).

Il SIP è la versione di *information package* gestita in fase di *immissione (ingest)*. Un determinato SIP è in generale il risultato di un accordo negoziato tra il Produttore e l'OAIS. Il concetto di SIP enfatizza il fatto che le risorse digitali fornite dal Produttore siano preparate e organizzate per essere gestite da OAIS. Per esempio, l'informazione conservata può essere l'aggregazione del contenuto fornito in molteplici SIP; o il Produttore può fornire l'informazione in un formato non supportato dall'OAIS (in questo caso si renderà necessaria una migrazione ad un altro formato); oppure i metadati forniti dal Produttore sono incompleti o inadeguati e occorre ampliarli durante il processo di immissione.

L'AIP è la versione di *information package* archiviata e conservata dall'OAIS. L'informazione archiviata e i relativi metadati associati rappresentano un singolo pacchetto logico all'interno di OAIS. Il modello non prescrive mai soluzioni tecnologiche determinate. Anche in questo caso le possibili soluzioni al rapporto tra oggetto digitale e metadati associati sono lasciate alla specifica implementazione: è possibile la completa integrazione fisica tra metadati e dati (p. es. un unico file con più "segmenti") o la distribuzione di dati e metadati in database separati ma logicamente in relazione.

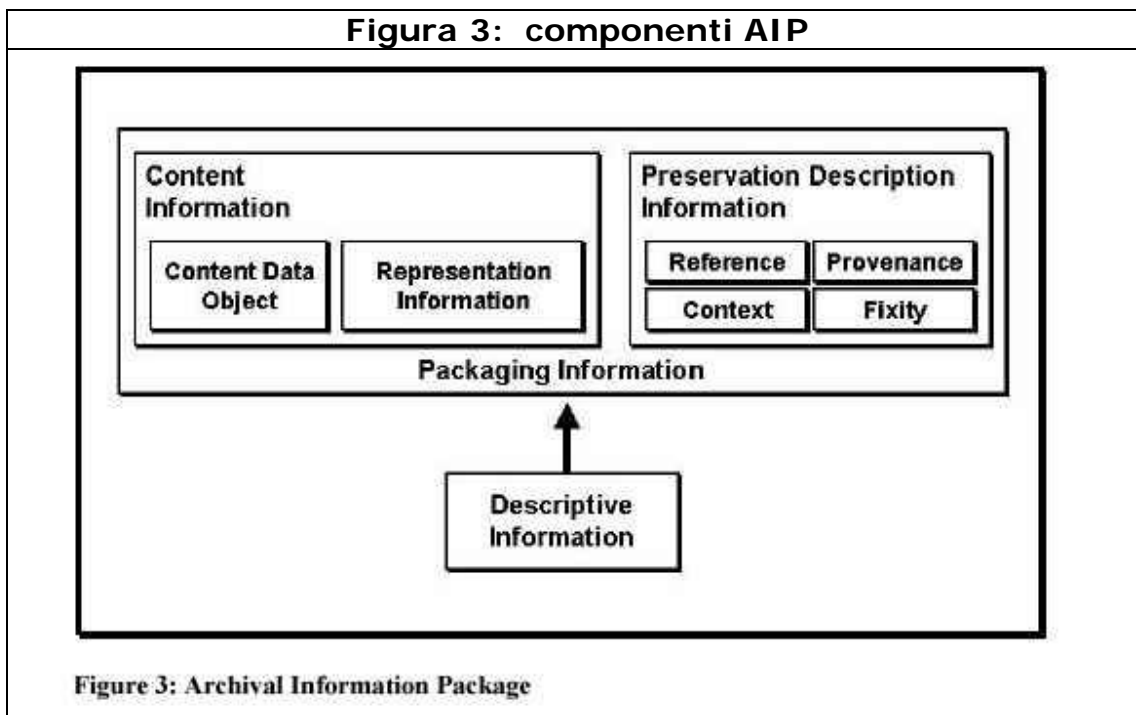
Infine il DIP, è la versione di *information package* consegnata all'utente in risposta ad una richiesta di accesso. Il concetto del DIP enfatizza il fatto che il pacchetto di informazioni diffuso dall'OAIS può essere differente per forma o contenuto da quello archiviato (AIP). Punti di differenziazione tra il DIP e AIP possono essere ad esempio:

- il formato (per esempio, un file può essere convertito dal TIFF al JPEG prima di essere inviato all'utente);
- i metadati forniti con il contenuto (è probabile che il DIP non contenga la serie completa dei metadati associati a un oggetto digitale archiviato, perché gran parte di questi possono non essere di interesse per un utente).



- L'AIP – l'Archival Information Package – è il perno della conservazione: è la versione di *information package* pensata per la conservazione nel lungo periodo. Data l'importanza dell'AIP per le responsabilità OAIS, è utile dare un'occhiata ravvicinata a questo *information package* ed esaminare le sue componenti chiave.

Come è stato detto un *information package* è formato da un insieme logico comprendente l'oggetto digitale (ciò che deve essere conservato) e i metadati necessari a garantire conservazione e accesso nel lungo periodo. Questo vale per SIP e DIP. Ma naturalmente vale soprattutto per AIP dato questo *information package* deve includere la serie completa dei metadati necessari a sostenere nel lungo periodo la conservazione del contenuto e la sua disponibilità per la "comunità di riferimento". Il *modello di riferimento* individua gli insiemi di metadati da includere in AIP.



La Figura 3 illustra le componenti d AIP. AIP si basa su un *Content Data Object* – li **dati** da conservare. Il *Content Data Object* può prendere la forma di una qualsiasi tipologia di materiale digitale: testo, immagini, video, database, programmi informatici, etc. Il *Content Data Object* può essere consistere di un singolo, autonomo file digitale – per esempio, un documento in formato PDF; può anche include diversi file, come un sito web che consiste di testo (file HTML) e immagini statiche

(file GIF o JPEG). OAIS come si è visto è responsabile sia della conservazione del *Content Data Object* a lungo termine, sia della sua disponibilità in una modalità che risulti essere "autonomamente" comprensibile dalla "comunità di riferimento".

Per la seconda responsabilità – rendere il *Content Data Object* disponibile in una modalità che risulti essere "autonomamente" comprensibile dalla "comunità di riferimento" – il *Content Data Object* deve essere accompagnato dalla *Representation Information* (informazione sulla rappresentazione o metadati): l'informazione necessaria a fruire e rendere comprensibile le sequenze di bit che costituiscono il *Content Data Object*. La *Representation Information* potrebbe includere la descrizione dell'ambiente hardware e software necessario a visualizzare il *Content Data Object* e/o ad accedere ai suoi contenuti, ma anche qualsiasi altro tipo di informazione necessaria. Per esempio, se il *Content Data Object* è un file ASCII contenente numeri, la *Representation Information* potrebbe indicare che il numeri corrispondono alla rilevazioni medie giornaliere della temperatura dell'aria a Manhattan, misurate in gradi Celsius, per il periodo 1972-2000.

E' stato prima ricordato che i confini della "comunità di riferimento" hanno un impatto sulla quantità dei metadati richiesti per sostenere il processo di conservazione. Si tratta appunto di metadati che costituiscono la *Representation Information*. In generale, più ampio è il confine della "comunità di riferimento", meno specializzata è la base di conoscenze della "comunità di riferimento". In questo caso occorrerà includere un maggior numero di metadati nella *Representation Information* per assicurare che l'informazione conservata resti fruibile e comprensibile alla "comunità di riferimento" nel lungo periodo.

Il **Content Data Object** e l'associata **Representation Information** sono conosciute insieme come **Content Information** – insieme di dati e metadati sufficienti ad assicurare la fruibilità e la comprensibilità nel lungo periodo per la "comunità di riferimento".

L'archiviazione di lungo periodo della *Content Information* (Informazione sul contenuto) richiede tuttavia ancora altri metadati necessari allo svolgimento dei processi di conservazione OAIS. Tra questi la *Preservation Description Information* (Informazione descrittiva per la conservazione), o PDI.

La PDI consiste di quattro componenti

1. La *Reference Information (Identificazione)* identifica in maniera univoca la *Content Information*. Esempi ne sono un identificatore interno generato dal sistema, o un ISBN.
2. La *Context Information (Contesto)* descrive le relazioni della *Content Information* con altre *Content Information* archiviate (p. es. versioni in formati diversi).
3. La *Provenance Information (Provenienza)* documenta:
  - la storia della *Content Information*, dalla sua creazione, a qualsiasi trasformazione subita
  - la catena di custodia;
  - qualsiasi intervento effettuato (p. es. l'aggiornamento dei supporti o la migrazione) per conservare la *Content Information* e i risultati di questo intervento.
4. La *Fixity Information (Autenticazione)* documenta i meccanismi messi in atto per assicurare l'integrità e l'autenticità della *Content Information* (impronta digitale MD5, firma digitale ecc)

**Content Information e Preservation Description Information** rappresentano insieme il contenuto digitale archiviato, i metadati necessari a fruirlo e a comprenderlo e quelli necessari per la conservazione.

La *Packaging Information (Informazione di impacchettamento)* è introdotta nel modello per raccogliere tutte queste componenti informative in un singolo pacchetto logico: in altre parole, la *Packaging Information* serve ad associare tutte le varie componenti di un AIP, permettendo loro di essere identificate e localizzate come una singola unità logica all'interno del sistema per l'archiviazione.

Infine, la *Descriptive Information (Informazione Descrittiva)* è l'informazione che permettere il recupero della *Content Information* da parte della "comunità di riferimento". Per esempio, la *Descriptive Information* potrebbe prendere la forma di un record in formato Dublin Core.

Le componenti sopra descritte – *Content Information (Content Data Object e Representation Information)*, *Preservation Description Information (Reference, Context, Provenance e Fixity Information)*, *Packaging Information*, formano insieme l'*Archival Information Package*, che a sua volta rappresenta la combinazione di informazione digitale conservata e una serie completa di metadati associati.