

1306  
28 NOV. 2013

ASL8

PG/2013/ 0133327 del 11/11/2013 ore 16,21

Mittente: POSTE ITALIANE SDA - GRANDI IMPRE

Assegnatario: Servizi Generali Alberghieri e

Classifica: 1.4.6 Fascicolo: 167 del 2013



Spett.le

ASL 8 CAGLIARI

Via Piero della Francesca 1

09047Cagliari

Roma, 3 ottobre 2013

**Oggetto: Contratto per il servizio GED Documenti - Proposta di modifica per pubblicazione Delibere e Determine.**

In riferimento al contratto per il servizio in oggetto, sottoscritto in data 05/06/2013, E seguendo le intese raggiunte, Vi comunichiamo la nostra proposta di modifica delle condizioni economiche come di seguito riportato:

- **Data-entry** degli elementi di indicizzazione univoca dei documenti in oggetto tramite interfaccia web della ASL;
- **Pubblicazione** dei documenti in oggetto su area dedicata del sito della ASL;

Le attività sopraindicate, quotate per i volumi complessivi di 4.000 documenti indipendentemente dalla tipologia documentale e dal numero di allegati, al prezzo di:

**€ 1,27 a documento(oltreIVA).**

La presente nota costituisce parte integrante e sostanziale del citato contratto.

Resta ferma ed impregiudicata ogni altra pattuizione prevista in contratto e non espressamente derogata con la presente nota.

Vogliate renderci copia della presente, riportata su Vs carta intestata, sottoscritta da un Vs Rappresentante munito di adeguati poteri in segno di integrale accettazione.

Cordiali saluti

Poste Italiane SPA – con socio unico  
Grandi Imprese e Pubbliche Amministrazioni

Valter Catoni  
Il Responsabile

Poste Italiane SpA – società con socio unico  
Sede Legale 00144 Roma - Viale Europa 190 Partita IVA 01114601006 Codice Fiscale 97103880585  
Capitale Sociale Euro 1.306.110.000 i.v. Registro delle Imprese di Roma n. 97103880585/1996

*Handwritten notes and signatures:*  
Dott. Ricci  
11/11/2013  
[Signature]

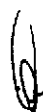
Il presente allegato è com-  
posto di n° 22 fogli  
di n° 22 pagine.

IL RESPONSABILE DEI SERVIZI  
GENERALI ALBERGHIERI E LOGISTICA

Ing. Raffaele Peratta

**ASL 8 CAGLIARI**  
**Allegato tecnico**

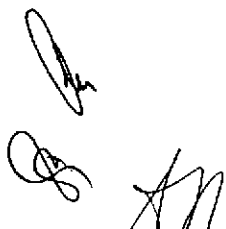
**Dematerializzazione della documentazione cartacea**  
**DELIBERE e DETERMINE**  
**con restituzione del cartaceo**



## Sommario

---

LA PIATTAFORMA GED DOCUMENT .....	3
ARCHITETTURA FISICA .....	3
ARCHITETTURA APPLICATIVA .....	4
LIVELLO PRESENTATION .....	4
LIVELLO BUSINESS .....	4
LIVELLO DATI .....	4
FUNZIONALITÀ OFFERTE .....	5
INTERFACCE UTENTE .....	5
INTERFACCIA WEB .....	5
WORKFLOW (GESTIONE DEL CICLO DI VITA) .....	6
CONTENT STORAGE SERVICES .....	7
VISUALIZZAZIONE DOCUMENTI .....	7
WEB SERVICES .....	8
GESTIONE DI PROCESSI DI BUSINESS .....	8
INTEGRAZIONI .....	9
INTEGRAZIONE CON LE PRINCIPALI APPLICAZIONI MICROSOFT E TERZE PARTI .....	9
GESTIONE DELLA SICUREZZA .....	10
GESTIONE ANOMALIE .....	10
PIATTAFORMA GED - CAPTIVA .....	11
FASI DEL PROCESSO DI DEMATERIALIZZAZIONE .....	11
ARCHITETTURA .....	12
SUPPORTO AVANZATO AL FAILOVER .....	13
DISTRIBUTED CAPTURE .....	13
SCANSIONE E ACQUISIZIONE MULTICANALE .....	13
CLASSIFICAZIONE INTELLIGENTE .....	14
VALIDAZIONE UTENTE .....	16
GESTIONE PROFILI E SECURITY .....	16
INDICIZZAZIONE E PUBBLICAZIONE DELIBERE E DETERMINE SUL SITO .....	16
PREMESSA .....	16
DESCRIZIONE PROCESSO .....	17



## LA PIATTAFORMA GED DOCUMENT

Il sistema GED di Poste Italiane si basa su un'infrastruttura tecnologia ad elevate prestazioni in grado di garantire i migliori standard di sicurezza e affidabilità con una completa aderenza alla normativa vigente.

Il sistema di Gestione Documentale presente all'interno del Data Center Poste Italiane è basato sulla piattaforma **EMC Documentum**.

Il sistema GED è fondato sul concetto di REPOSITORY: tutti i documenti e le immagini archiviate possono essere mantenute all'interno di un apparato di storage, il cui accesso è consentito al Cliente tramite interfaccia WEB con possibilità di ricerca, consultazione e gestione dei contenuti archiviati; contestualmente il servizio di archiviazione potrà essere richiesto attraverso l'inoltro via telematica dei documenti archiviati, o la creazione di supporti ottici CD/DVD completi del software di consultazione.

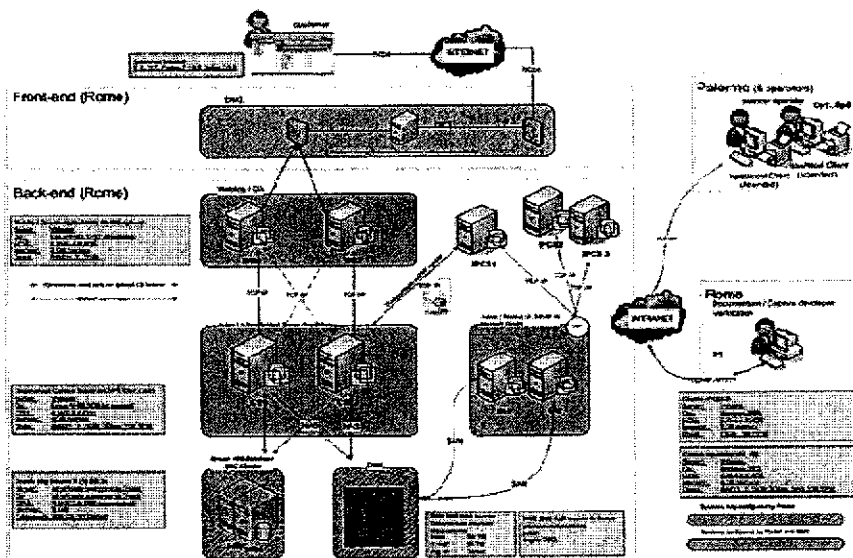
## ARCHITETTURA FISICA

L'architettura fisica del sistema GED è la seguente:

- Reverse Proxy IBM http Server 6.1, Apache web server (RP1);
- Data Server Oracle 10 G in alta affidabilità (PB1, PB2);
- Content Server con SO Red Hat Enterprise Linux 5.0 (CS1, CS2);
- Application Server con SO Red Hat Enterprise Linux 5.0 e Web Server IBM WS 6 (WS1, WS2);
- Storage dati di tipo SAN / NAS (EMC DMX), EMC Centera;
- Client Acquisizione con SO Windows 2003 (OP1);
- Image Processing Component Server con SO Windows 2003 (IPCS1, IPCS2).

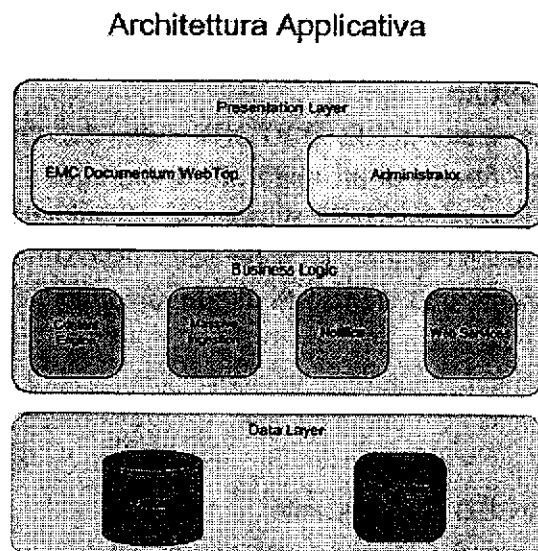
Di seguito si mostra uno schema grafico:

FIGURA 4 – SCHEMA GRAFICO DELL'ARCHITETTURA FISICA



## ARCHITETTURA APPLICATIVA

L'architettura applicativa del sistema GED è sviluppata sul classico pattern MVC a 3 livelli. Di seguito si mostra uno schema grafico:



### **LIVELLO PRESENTATION**

Il livello **presentation** è sviluppato con interfacce web fruibili da browser.

Nello specifico si ha:

- Modulo utente - portale: gli utenti "clienti" del sistema raggiungono l'applicazione attraverso pagine e form HTML. L'applicazione principale si basa sul modulo software EMC-Documentum WebTop, mediante il quale possono essere costruite web application ad-hoc per i clienti con funzionalità e layout personalizzati;
- Modulo utente – web services: una serie di web services espongono le principali funzionalità;
- Modulo amministrazione: l'amministratore di sistema fruisce dell'applicazione attraverso pagine e form HTML, basate sul modulo software EMC-Documentum (DA- Documentum Administrator).

### **LIVELLO BUSINESS**

A livello **business** il sistema GED si presenta nel seguente modo:

- Una serie di moduli software in linguaggio Java (J2EE) che costituiscono le applicazioni verticali e le personalizzazioni della piattaforma EMC-Documentum;
- Content Engine (document management e process management);
- Massive Ingestion (caricamento e indicizzazione documenti);
- Notifica (notifica completamento processo di archiviazione).

### **LIVELLO DATI**

A livello dati il sistema utilizza:

- RDBMS Oracle per la memorizzazione dei metadati e dei dati di configurazione;
- Storage di tipo NAS / SAN (EMC NSX e DMX), e EMC Centera per lo stockaggio sicuro dei documenti.

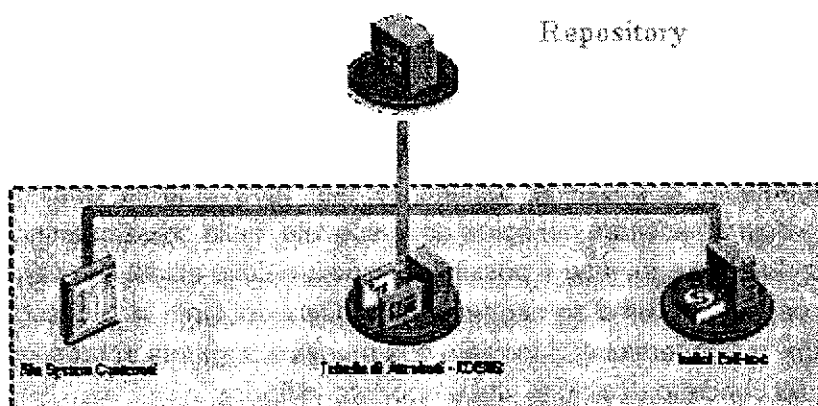
I contenuti digitali gestiti da Documentum sono memorizzati all'interno del cosiddetto repository. Il repository è un archivio che permette la memorizzare e la gestione di oggetti.

Dal punto di vista fisico il repository è implementato coordinando 3 fondamentali servizi:

- uno o più file system, per la memorizzazione dei contenuti non strutturati (immagini, documenti PDF, etc);
- uno ed un solo database, per la memorizzazione di tutte le informazioni strutturate; tali informazioni, i metadati, sono in generale associati ai contenuti non strutturati per facilitare il loro reperimento durante una ricerca, implementare la sicurezza durante il loro accesso, implementare processi di business;
- opzionalmente uno o più fulltext index server per estendere la ricerca dei documenti a partire dal loro contenuto e/o estendere le funzionalità di ricerca dei metadati rispetto ai meccanismi di ricerca base, esposti indirettamente dal database.

Il repository oltre a memorizzare oggetti relativi a contenuti gestiti direttamente dall'utente finale, permette la gestione di tutte le informazioni necessarie al suo corretto funzionamento, come ad esempio utenti, gruppi, permessi di accesso associati agli oggetti (le cosiddette ACL, Access Control List).

FIGURA 6 – SCHEMA GRAFICO LIVELLO DATI



## FUNZIONALITÀ OFFERTE

La piattaforma GED di Poste Italiane include un tool per il disegno di interfacce utente WEB, altamente configurabile e ottimizzata per l'elaborazione di operazioni di alto volume, il recupero dei documenti e la visualizzazione dei documenti in funzione del ruolo o del contesto. L'interfaccia utente è potenziata grazie alla disponibilità di un visualizzatore integrato, all'accesso al repository e ai servizi di gestione del contenuto.

## INTERFACCE UTENTE

La piattaforma GED di Poste Print, mette a disposizione diversi tipi di interfacce per la consultazione e la gestione del sistema documentale.

## INTERFACCIA WEB

L'interfaccia principale è un applicazione basata sul Web che consente agli utenti finali di accedere alla gestione dei documenti. Vi si accede attraverso un browser, pertanto non è necessario software aggiuntivo pre-installato sulla macchina dell'utente.

L'interfaccia presenta le seguenti caratteristiche di interesse:

- Accesso ai contenuti, includendo controlli di sicurezza, controllo di versioni, cicli di vita, ricerche, ecc.;
- Gestione avanzata dei contenuti, con workflow, documenti virtuali, sottoscrizione a servizi di notifica;
- E' disponibile in diverse lingue;
- Offre una gestione semplificata delle Ricerche attraverso uno strumento di ricerca in ogni pagina visualizzata;
- Facilità e rapidità di configurazione.

La tecnologia presente implementa interfacce role-based. Grazie a questa caratteristica l'interfaccia e le funzionalità presenti nei client, sono in grado di modificarsi dinamicamente a seconda del ruolo dell'utente che vi accede.

Alcuni utenti, ad esempio, potrebbero vedere un'interfaccia fortemente semplificata dove siano visibili solo i componenti per la ricerca e l'organizzazione per categorie.

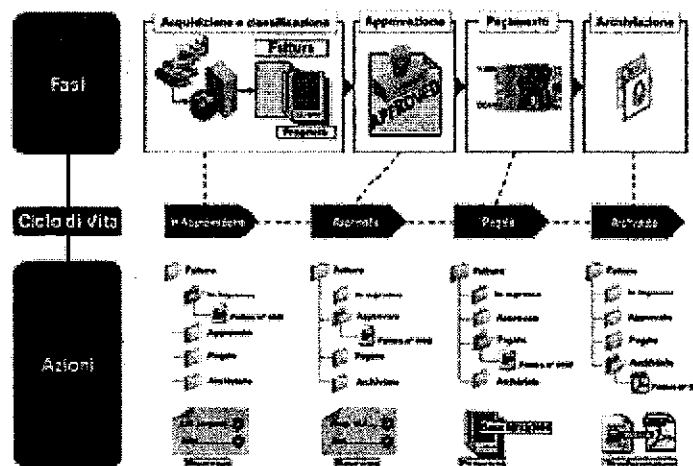
I responsabili invece, o gli operatori che necessitano di più strumenti, come l'inbox per avviare e partecipare ai workflow, la struttura gerarchica per navigare all'interno delle cartelle, le funzionalità di versioning e così via, possono essere configurati con un profilo che permette di visualizzare un'interfaccia avanzata.

## WORKFLOW (GESTIONE DEL CICLO DI VITA)

Grazie a questa tecnologia è possibile definire le fasi attraverso le quali ogni tipologia di documento/contenuto dovrà transitare nel corso della sua vita, dalla creazione fino all'archiviazione e all'eventuale distruzione. Ad ogni fase è possibile associare delle azioni che verranno eseguite in maniera trasparente dal server come ad esempio la modifica dei diritti di accesso, lo spostamento da uno storage all'altro, la conversione di formato o la modifica aggiunta di proprietà specifiche.

Un'esemplificazione del concetto di ciclo di vita è riportato in figura.

FIGURA 7 – ESEMPIO DI WORKFLOW DOCUMENTALE



In generale, è importante che la definizione e la gestione delle fasi attraverso le quali un qualunque documento dovrà passare sia completamente separata ed indipendente da eventuali processi di business. Ciclo di vita e processi di business, possono al più interagire tra loro ma rimangono comunque due cose fondamentalmente diverse.

Non tutti i documenti devono essere gestiti da processi automatizzati. Possono esistere molte categorie di documenti da gestire e molti contesti nei quali l'automazione tramite workflow non è richiesta o non è implementabile. Ciò nonostante, questi documenti avranno comunque un loro ciclo di vita che non solo permetterà di definire una per una tutte le fasi che dovranno attraversare ma consentirà di configurare impostazioni fondamentali come la sicurezza, la distribuzione del documento all'interno dell'organizzazione gerarchica dei folder ed eventualmente anche l'esecuzione di procedure e metodi personalizzati.

In questo modo per esempio, si potrà impostare dinamicamente la sicurezza di una determinata tipologia di documenti in modo tale che in fase di approvazione essi saranno visibili solo dal gruppo di utenti che li deve processare e, una volta approvati, essi verranno non solo archiviati in un folder particolare ma saranno accessibili da un altro gruppo (o utente) particolare.

Una volta quindi definiti gli stati attraverso i quali la tipologia di documento in esame dovrà transitare nel corso della sua esistenza, sarà possibile per ciascuno di essi, assegnare delle azioni che verranno eseguite in modo sicuro e trasparente dal sistema GED.

Grazie al ciclo di vita, non solo sarà possibile definire gli stati che potrà assumere un documento ma anche definire tutte le azioni che dovranno essere eseguite ad ogni cambio di stato, come modifica di autorizzazione di accesso, modifica delle proprietà del documento, modifica delle azioni personalizzate che possono essere eseguite sul documento. GED offre uno strumento grafico specifico per la definizione del ciclo di vita. Grazie a questo strumento è possibile definire e configurare i vari elementi che costituiscono il ciclo di vita, come gli stati, le regole di passaggio da uno stato al successivo, le azioni che il server dovrà eseguire ad ogni stato. In figura è mostrato un esempio di configurazione.

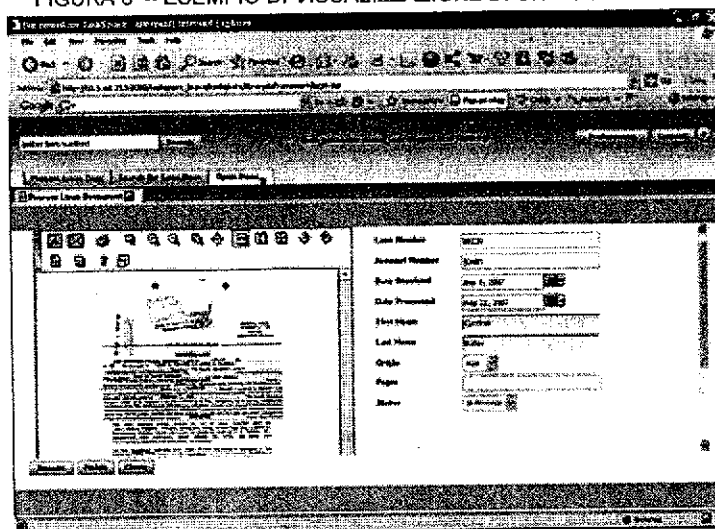
## CONTENT STORAGE SERVICES

Gestione intelligente dell'infrastruttura di storage. Grazie alla gestione completa del ciclo di vita la piattaforma sarà in grado di determinare in base a policy preimpostate, l'area di storage più adeguata per la fase di vita in cui si trovano i documenti.

## VISUALIZZAZIONE DOCUMENTI

L'interfaccia offre un visualizzatore dei documenti che permette all'utente di visualizzarli senza l'ausilio di applicazione o visualizzatori esterni.

FIGURA 8 – ESEMPIO DI VISUALIZZAZIONE DI UN DOCUMENTO





Per una visualizzazione contestuale dei documenti grazie all'interfaccia role-based non sarà necessario fornire due diversi client personalizzati con conseguente dispendio di risorse e tempo per la loro realizzazione ed il loro mantenimento.

I dati e i documenti possono essere estratti per accedere ed elaborare le richieste dei clienti. Le informazioni possono essere recuperate mediante le funzionalità di ricerca dei metadati e full-text sfogliando le cartelle, i fascicoli e le categorie nel repository.

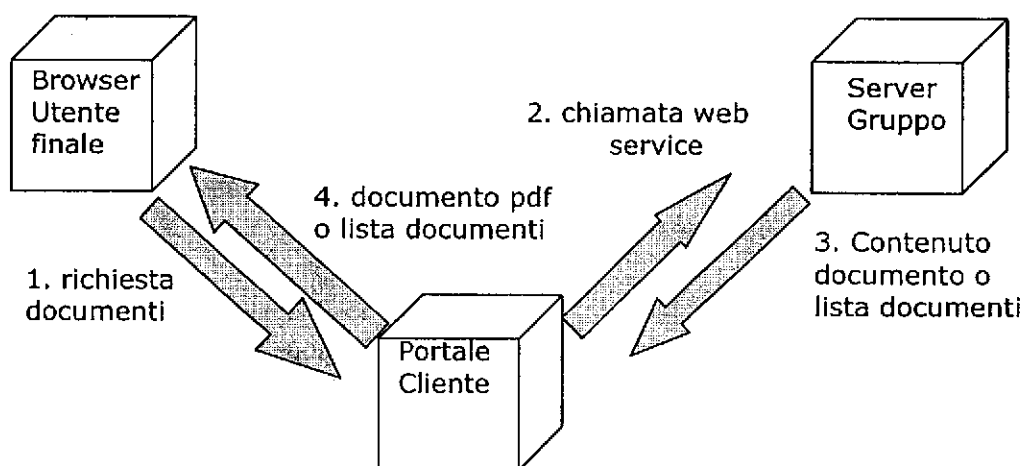
La creazione di 'fascicoli virtuali' permette la ricerca e la consultazione contestuale dei documenti archiviati e delle immagini di ritorno corrispondenti, con tutte le informazioni associate ai diversi oggetti. A titolo di esempio questa funzionalità potrà essere utilizzata per la creazione di documenti virtuali che contengano per ogni comunicazione raccomandata, la relativa immagine della cartolina e la busta inusitata con le informazioni associate.

## **WEB SERVICES**

L'interfacciamento con il servizio avviene attraverso web service.

La modalità di comunicazione "Web service" fornisce le funzionalità di ricerca dei documenti attraverso un insieme di operazioni invocabili via web.

FIGURA 9 – SCHEMA DI FUNZIONAMENTO WEB SERVICES



Quando il portale del Cliente riceve la richiesta di uno o più documenti da parte dell'utente finale utilizza i metodi disponibili nel web service per ottenere il documento o la lista dei documenti cercati. Una volta ottenuto il documento o la lista dei documenti è il portale stesso che si occupa di restituirli all'utente finale.

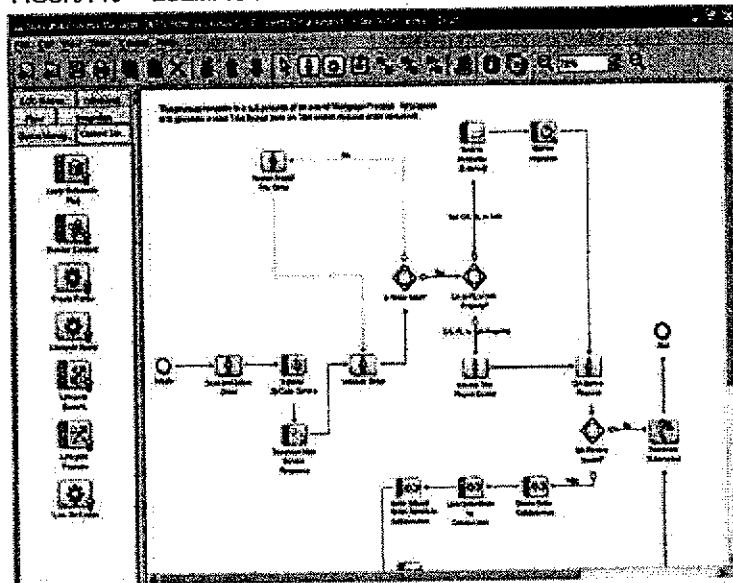
Questo schema di interazione consente di avere il pieno controllo della comunicazione con l'utente finale in quanto quest'ultimo comunica sempre e solo con il portale del Cliente. Eventuali errori nelle ricerche dei documenti vengono riportati al portale del Cliente che può implementare la politica della gestione dell'errore più opportuna.

## **GESTIONE DI PROCESSI DI BUSINESS**

L'interfaccia include anche la visualizzazione logica del processo secondo un modello di visualizzazione organizzato in scheda.

Una volta disegnato il processo utilizzando l'apposito strumento grafico (Process Builder), le schede permettono di contestualizzare la visualizzazione dei documenti, la ricerca, l'elenco delle pratiche su cui lavorare e il rispettivo stato di lavorazione, in funzione del processo stesso.

FIGURA 10 – ESEMPIO DI VISUALIZZAZIONE LOGICA DEL PROCESSO



## INTEGRAZIONI

E' possibile integrare il sistema con sistemi esterni, come:

- ECM disponibili sul mercato;
- Database esterni compatibili ODBC;
- Esportazione dei dati su File System;
- SAP, mediante un'integrazione opzionale;
- Sharepoint, mediante un'integrazione opzionale.

E' inoltre possibile effettuare sviluppi custom che estendano le funzionalità già esistenti del sistema.

## INTEGRAZIONE CON LE PRINCIPALI APPLICAZIONI MICROSOFT E TERZE PARTI

La piattaforma GED può essere nativamente integrata con numerose applicazioni client per la creazione e la modifica dei documenti, con il vantaggio di rendere l'esperienza utente praticamente trasparente all'infrastruttura sottostante.

L'integrazione è garantita dal modulo Documentum Applications Connector, che permette alle varie applicazioni di comunicare con il repository attraverso voci di menù dell'interfaccia nativa dell'applicazione.

Il connettore è basato sul layer di integrazione di Documentum denominato WDK che mette a disposizione dei componenti web pronti all'uso che costituiscono tra l'altro, l'infrastruttura di base di tutte le applicazioni client web di Documentum come Documentum Webtop e Documentum Administrator.

*[Handwritten signatures and initials]*

Questa caratteristica consente quindi di integrare le applicazioni client direttamente con il presentation layer di Documentum senza essere obbligati ad installare applicazioni impegnative sia in termini di spazio su disco che di configurazione su ogni client che necessita l'integrazione.

Le funzionalità offerte nativamente dall'integrazione includono:

- Creazione di documenti basati su template salvati nel repository;
- Ricerca per proprietà e full-text di documenti;
- Visualizzazione e gestione delle proprietà (metadati) di ciascun documento;
- Gestione del versioning;
- Gestione del Ciclo di Vita dei documenti salvati;
- Menù dinamici sensibili al contesto.

Queste funzionalità sono esposte in un menù che si aggiunge a quelli già presenti nelle varie applicazioni.

Per i clienti che hanno necessità di forti configurazioni o l'aggiunta di funzionalità specifiche sarà comunque possibile rilasciare una specifica web application personalizzata.

## **GESTIONE DELLA SICUREZZA**

I contenuti del sistema GED sono completamente protetti: ciascun elemento è associato a un elenco di controllo degli accessi che definisce quali utenti, gruppi o ruoli possono accedere a esso e quali operazioni possono eseguire.

Ogni utente sarà identificato da uno username e una password, e gli sarà richiesto di inserirli al momento dell'autenticazione.

GED può autenticare direttamente gli utenti, mantenendo traccia della coppia valori username / password oppure può delegare l'autenticazione (e quindi la conservazione e gestione delle password) a sistemi esterni come il sistema operativo sul quale è installato il Content Server oppure a uno o più server LDAP. Il sistema permette di utilizzare queste diverse modalità di autenticazione contemporaneamente all'interno dello stesso repository.

GED permette inoltre di gestire sofisticati meccanismi di auditing che consentono agli amministratori di conoscere chi è entrato nel sistema e quali azioni ciascun utente ha compiuto. E' possibile compiere un monitoraggio su un set di eventi molto esteso, tra i quali quelli legati alle seguenti attività:

- inserimento nel sistema di un nuovo documento di una data tipologia;
- eventi specifici, indipendentemente dagli oggetti su cui essi avvengono;
- eventi relativi a specifici cicli di vita dei documenti.

Avere accesso ad uno specifico documento non significa avere automaticamente accesso a tutti gli oggetti che si trovino nello stesso "contenitore". La granularità degli accessi può essere molto specifica.

Utenti o gruppi di utenti diversi possono lavorare in parallelo su aree diverse della stessa struttura di un documento, il tutto sempre garantendo la sicurezza o la privacy per ogni singolo documento facente parte della struttura.

## **GESTIONE ANOMALIE**

Per un completo controllo di tutti gli aspetti della fase di archiviazione dei documenti nella piattaforma GED, si focalizza l'attenzione sulle seguenti tipologie di emergenza:

- **Infrastrutturale:** si considerano tutti gli eventuali problemi relativi a macchine, comunicazione, aree di storage e DB;

- **Applicativa:** le problematiche che si è pronti a gestire riguardano tutto il middleware a supporto della GED di Poste Italiane e gli applicativi che compongono il servizio;
- **Esercizio:** a questo livello si gestiscono tutte le problematiche relative alla ricezione dei flussi, all'importazione dei singoli documenti e alla erogazione del servizio di ricerca ed esposizione.

Per evidenziare prontamente gli eventuali problemi, si adottano sia soluzioni di mercato sia applicativi appositamente sviluppati; in questo modo si dispone dello strumento più adatto ad individuare e segnalare la problematica specifica.

La piattaforma applicativa utilizzata è HP OpenView, che consente di effettuare monitoraggi a livello infrastrutturale ed applicativo, evidenziando problematiche su hardware, connessioni di rete, applicativi, riempimenti delle aree di storage e stato dei DB.

In particolare, consente di attivare allarmi al verificarsi di precise condizioni e di allertare i responsabili dei vari sistemi con mail, SMS e segnalazioni sui cruscotti di controllo. Talvolta gli allarmi prevedono opportune azioni automatiche atte alla chiusura del problema.

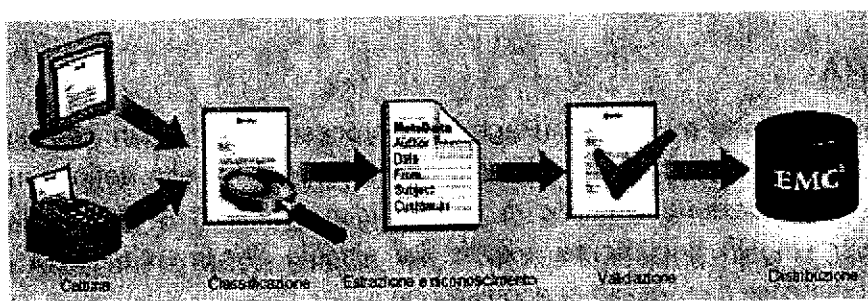
Si è inoltre deciso di dotare il gruppo di supporto, di tool custom che consentono di monitorare lo stato dei carichi dei flussi in ingresso e l'effettiva disponibilità del servizio offerto.

Una volta rilevato e segnalato il problema, il gruppo di supporto GED si occupa di analizzare in dettaglio tutte le informazioni a disposizione. La fase successiva consiste nella risoluzione della problematica, o nello smistamento verso un reparto specializzato (supporto di rete, supporto DB etc...).

## **PIATTAFORMA GED - CAPTIVA**

Il processo di dematerializzazione di Poste Italiane è composto dalle fasi di cattura, identificazione, estrazione delle informazioni, validazione e distribuzione; la piattaforma software utilizzata per tale processo è basata sul software EMC CAPTIVA, la cui suite di moduli è direttamente integrata nella piattaforma EMC DOCUMENTUM, presente nel sistema GED di Poste Italiane.

FIGURA 11 – SCHEMA LOGICO DI FUNZIONAMENTO DELLA PIATTAFORMA



### **FASI DEL PROCESSO DI DEMATERIALIZZAZIONE**

La fase di cattura opera in modalità multi-canale interfacciando scanner, fax, web client, file system ed altri tipi di sorgenti. Sono supportati i più diffusi formati di tipo immagine, anche in modalità multi pagina.

Le immagini possono essere elaborate automaticamente attraverso filtri (MR\_RF\_25) tra i quali spiccano quelli di rimozione dei disturbi, di raddrizzamento e di sottrazione (del colore, del bordo, ecc.).

Tipicamente il processo di dematerializzazione del cartaceo, richiede la scansione differita e per questo la soluzione mette a disposizione diverse tecniche quali la rilevazione dei codici a barre, la patch-code, la blank page.

Il processo continua con la fase di classificazione delle immagini catturate secondo tecniche che consentono di identificare e analizzare pattern grafici o testuali attraverso tecniche matematiche o di Intelligent Character Recognition (ICR):

- il confronto con modelli predefiniti del documento stesso o di sue porzioni;
- il confronto con documenti o porzioni di documenti simili già classificati;
- l'individuazione di zone distintive del documento;
- criteri posizionali e di formato;
- altri algoritmi matematici.

Le diverse tecniche possono anche essere utilizzate in serie per aumentare la percentuale di successo del riconoscimento automatico ed è possibile gestire anche i documenti scritti a mano.

In seguito alla classificazione, nella fase di estrazione e riconoscimento vengono applicate tecniche di Character Recognition avanzato (OCR, ICR, OMR) per estrarre le informazioni di contenuto sia in maniera fissa (OCR zonale) che dinamica (FreeForm). Analogamente alla fase precedente sono disponibili diverse tecniche eventualmente applicabili in serie.

Per ogni informazione estratta è possibile impostare una serie di attributi quali l'obbligo o meno di validazione, la visibilità, la conversione in un formato specifici in termini di lunghezza, presenza di caratteri e numeri, ed altro ancora.

A questo punto, se necessario, nella fase di validazione viene richiesto l'intervento umano per controllare e correggere eventuali eccezioni verificatesi nella fase di estrazione e riconoscimento. L'interfaccia utente è intuitiva e l'operatore ha immediatamente evidenza sia della struttura del documento che del tipo di intervento richiesto.

L'operatore può anche effettuare un OCR esplicito selezionando con il mouse la zona sulla pagina e guidando manualmente il riconoscimento richiedendo anche la validazione dei valori estratti direttamente alle applicazioni di back-end.

Il punto di forza della piattaforma risiede nella sua modularità (possibilità di estendere la soluzione), legata alla sua capacità di supportare esigenze di alti volumi.

Di seguito vengono delineate le caratteristiche principali della piattaforma.

## **ARCHITETTURA**

- Capacità di elaborazione multilivello (Page-level processing vs. batch processing). In questo modo ogni modulo del processo di acquisizione lavorerà al livello definito in fase di design (tipicamente classificazione a livello di batch, riconoscimento e OCR a livello di pagina, validazione per singolo documento, export per singolo documento). Sarà quindi possibile modellare nel dettaglio il processo di acquisizione ;
- Scalabilità dei moduli in orizzontale/verticale e plug-in di nuovi moduli (aggiunta di un modulo che espleta nuove funzionalità o incremento di quelli esistenti per aumentare la capacità di elaborare in parallelo): qualora il volumi crescano o sia necessario gestire particolari picchi, sarà possibile mettere in parallelo i moduli che richiedono livelli di elaborazione superiore.

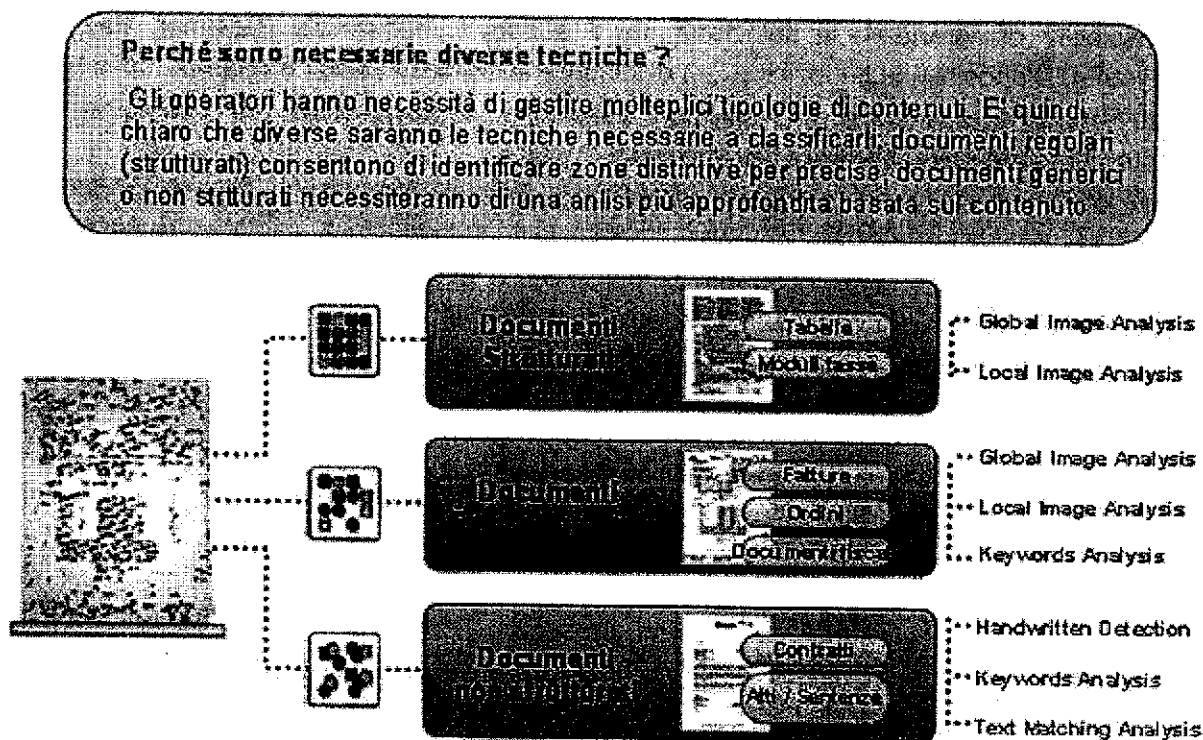
FIGURA 12 – SCHEMA LOGICO DI FUNZIONAMENTO



## CLASSIFICAZIONE INTELLIGENTE

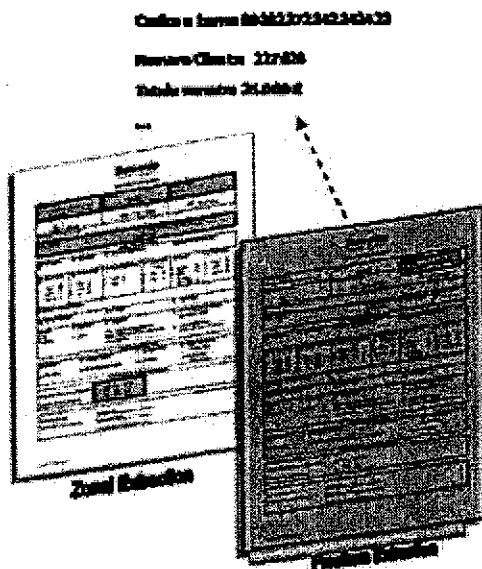
- Global Image Analysis: basata su un approccio che analizza gli elementi strutturali della pagina; nella fase di setup è previsto la funzionalità di autoapprendimento attraverso cui il sistema prende in pasto migliaia di documenti e crea in maniera automatica dei modelli di classificazione attraverso algoritmi matematici;
- High Precision Anchor (HPA): definito attraverso l'identificazione di zone distintive dell'immagine; tecnica semplice e facile da configurare in cui vengono definite delle zone distintive per il template;
- Keyword: viene riconosciuto il testo e vengono applicate regole posizionali e di formato (documento non strutturati ma riconoscibili da termini specifici); vengono cercate parole in base a punti di riferimento, posizioni, presenza di altro testo e secondo un certo formato specificabile attraverso la sintassi delle Espressioni Regolari;
- Global Text Matching: è una tecnica unica e sofisticata per la classificazione di documenti con formato differente ma che hanno contenuto di testo simile (documento non strutturati); vengono riconosciute e confrontate le frasi del documento;
- Handwritten: identificazione dei documenti scritti a mano.

FIGURA 14 – ESEMPIO DI TECNICHE DI CLASSIFICAZIONE INTELLIGENTE DI DOCUMENTI

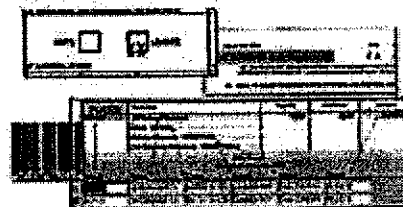


- Tecniche avanzate di estrazione delle informazioni (Character Recognition);

FIGURA 15 – ESEMPIO DI TECNICHE DI ESTRAZIONE DELLE INFORMAZIONI

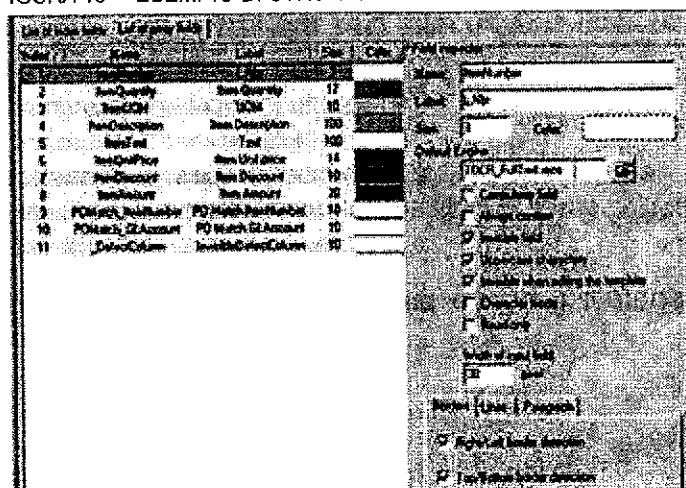


- Supporto a motore OCR OCR in stile griglia
- Applicazione di OCR su più pagine (comparsa pagina)
- Voting dei diversi motori (applicazione di più OCR su ogni simbolo da validare)
- Estrazione personalizzata 1D e 2D
- Configurazione e test immediati di ogni "template" OCR
- Lettura di diversi documenti in un unico template OCR
- Abilitazione e test di ogni motore OCR
- Abilitazione e test di ogni motore OCR
- Configurazione e test immediati di ogni "template" OCR
- Configurazione e test immediati di ogni motore OCR
- Configurazione e test immediati di ogni motore OCR



- Implementazione della tecnica del Voting (utilizzo di più motori contemporaneamente sullo stesso indice/campo);
- Possibilità di specificare il motore di OCR specialistico su ogni campo presente nella pagina;
- OCR Zonale o FreeForm;
- motore specializzato per l'estrazione delle righe da tabelle (es. righe di fatture, ordini, ecc.
- Configurabilità della parte di classificazione. L'utente ha a disposizione uno strumento di facile utilizzo per la configurazione e test immediata sia dei modelli di validazione che dei motori di estrazione (indicizzazione dati). Sarà quindi possibile integrare la soluzione con nuove tipologie di documenti senza modifiche al flusso di caricamento e sviluppo di codice complesso. Per ogni indice/campo è possibile definire graficamente una serie di parametri quali: obbligo o meno di validazione; Visibilità; Formato del campo (lunghezza, presenza caratteri, presenza numeri, ....), Motore OCR con cui viene fatta l'estrazione, con possibilità di selezionare uno specifico motore per ognuno dei diversi indici della pagina; Filtro grafico sulla zona oggetto di OCR di ogni campo (sottrazione del fondo nel caso di moduli a caselle, eliminazione del disturbo del disturbo, ecc.); ogni campo/indice può quindi essere ottimizzato in maniera specifica; nelle due figure seguenti è possibile avere una vista degli strumenti grafici di configurazione.

FIGURA 16 – ESEMPIO DI STRUMENTI GRAFICI DI CONFIGURAZIONE

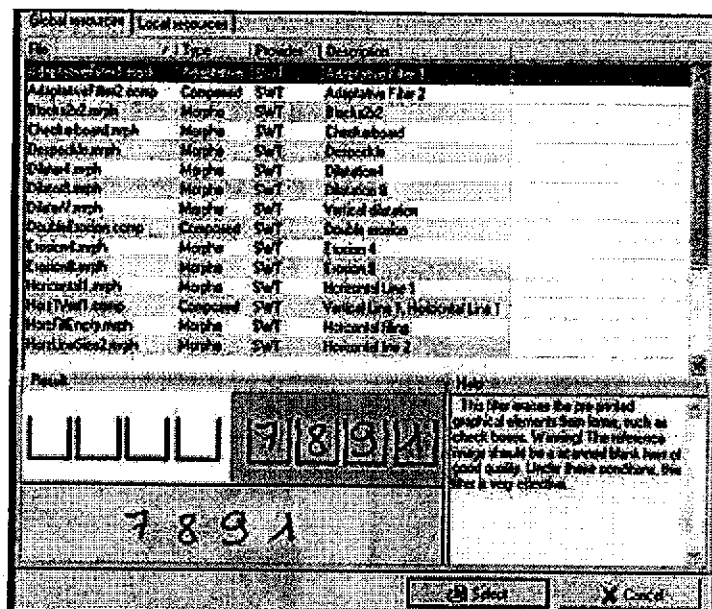


*[Handwritten signature]*

*[Handwritten signature]*

*[Handwritten signature]*





## VALIDAZIONE UTENTE

- Possibilità di validare su Database;
- Configurazione di regole cross pagina (es. controllo tra l'ammontare espresso su una fattura e la somma dei valori sugli allegati);
- Rubber band (OCR al volo);
- Alberatura intuitiva e evidenza dell'intervento richiesto;
- Possibilità di archiviare direttamente nel sistema di Gestione Elettronica Documentale quei documenti/indici che non necessitano di validazione (disponibilità immediata dei documenti "pronti", senza attendere la completa elaborazione del lotto).

## GESTIONE PROFILI E SECURITY

- Gestione dei ruoli e degli utenti avanzata;
- Supporto per Windows Encrypted File System (EFS);
- Possibilità re-indirizzamenti dipartimentali.

## INDICIZZAZIONE E PUBBLICAZIONE DELIBERE E DETERMINE SUL SITO

### **PREMESSA**

Il cliente ha espresso la necessità, ad integrazione del processo di dematerializzazione, di poter effettuare l'attività di indicizzazione e pubblicazione delle delibere e delle determine sul proprio sito web.

Tale attività partono dall'output generato dal processo di dematerializzazione, che presenta il seguente formato di file:

Formato file di output delibera:

a) Delibera: DEL\_NRDELIBERA\_GGMMAAAA\_TITOLODELIBERA.PDF

- ( documento principale)
- b) Per ogni allegato della delibera: DEL\_NRDELIBERA\_GGMMAAAA\_Ann.PDF  
(non progressivo allegato)
- c) Indici : DEL\_NRDELIBERA\_GGMMAAAA.XML

Formato file di output determina:

- a. Determina: DET\_RDETERMINA\_GGMMAAAA\_OGGETTODETERMINA.PDF  
( documento principale)
- b. Per ogni allegato della delibera: DET\_NRDELIBERA\_GGMMAAAA\_Ann.PDF (nn  
progressivo allegato)
- c. Indici : DET\_RDETERMINA\_GGMMAAAA.XML

## DESCRIZIONE PROCESSO

A seguire del processo di dematerializzazione, descritto in dettaglio nei paragrafi di cui sopra, è stato richiesto un'ulteriore attività che viene svolta presso l'isola digitale di Cagliari; tale processo integrativo si suddivide in due macro aree di intervento:

- indicizzazione dei documenti pdf prodotti nel processo di scansione;
- pubblicazione dei documenti pdf sul sito della ASL.

In dettaglio l'operatore, per ogni documento da processare effettua i seguenti passaggi:

- ricerca del documento pdf dell'atto e del documento pdf degli eventuali allegati, residenti su server FTP di Poste Italiane (isola digitale), sulla base dei campi indice utilizzati nella fase di dematerializzazione;
- download del file pdf così individuato e salvataggio del documento pdf presso il client locale dell'isola digitale;
- data entry attraverso l'interfaccia web messa a disposizione dall'ASL popolando i seguenti campi:
  - TIPO ATTO: è prevista una combo box, che prevede la selezione delle due tipologie documentali presenti a sistema (delibere o determine);
  - TIPO DOCUMENTO: Altro, Bando, Concorso, Gara, Procedura
  - STATO: Pubblicato, Bozza
  - TITOLO: è previsto un campo per l'inserimento del titolo (ad esempio: "Delibera n. \_\_ del \_\_\_\_")
  - DATA SEDUTA
  - DATA PUBBLICAZIONE
  - DATA E ORA SCADENZA:
  - DESCRIZIONE: è previsto un campo per l'inserimento dell'oggetto dell'atto da pubblicare (ad esempio: "Gara a procedura aperta per l'affidamento della fornitura, in due lotti, di autoambulanze di soccorso per i Presidi Ospedalieri dell'Azienda e per il Servizio di Emergenza Sanitaria 118. Aggiudicazione in favore della Società XXXXXXXX")



- upload dei documenti pdf dell'atto e degli eventuali allegati (frame ALLEGATI, funzionalità SFOGLIA, tasto INSERISCI) sul server del cliente a completamento del processo di indicizzazione;
- selezione dei documenti pdf precedentemente caricati sul server del cliente (frame ALLEGATI, funzionalità SFOGLIA, tasto SELEZIONA), in tal modo viene creato il codice html che renderà il documento disponibile per la consultazione;
- pubblicazione dell'atto sul server della ASL (tasto INSERISCI sotto il frame ALLEGATI).

Ultimata tale procedura il pdf viene reso disponibile sul server della ASL per la consultazione definitiva.

Si riportano a seguire le schermate del sito web in cui viene visualizzata l'attività eseguita dall'operatore con il fine di pubblicare i documenti pdf scansionati ed indicizzati e renderli quindi disponibili alla consultazione on line.

Dopo il login dell'utente, l'interfaccia web presenta la schermata della funzionalità di ricerca; le maschere per il data entry e l'upload dei documenti pdf vengono presentate dopo aver cliccato su INSERISCI NUOVA VOCE.



*[Handwritten signature]*

*[Handwritten signature]*

File Modifica Visualizza Preferiti Strumenti

Precedente Successivo Opzioni

Atti per aslcagliari

Tipo atto \*

Tipo documento ☒ Altro ☐ Banda ☐ Concorsi ☐ Gara ☐ Procedura

Stato ☒ Pubblicato ☐ Bozza

Titolo \*

Data seduta

Data pubblicazione \* 20/06/2013

Data e ora scadenza

Data rettifica

Altro soggetto

Descrizione \*

Allegati

[Ritorna alla form di ricerca](#)

File Modifica Visualizza Preferiti Strumenti

Precedente Successivo Opzioni

Ricerca documenti per aslcagliari

Inserire il testo da cercare:

Operatore: ☒ and ☐ or Tipo controllo:

Visibile dal web ☒ TUTTI ☐ SI ☐ NO

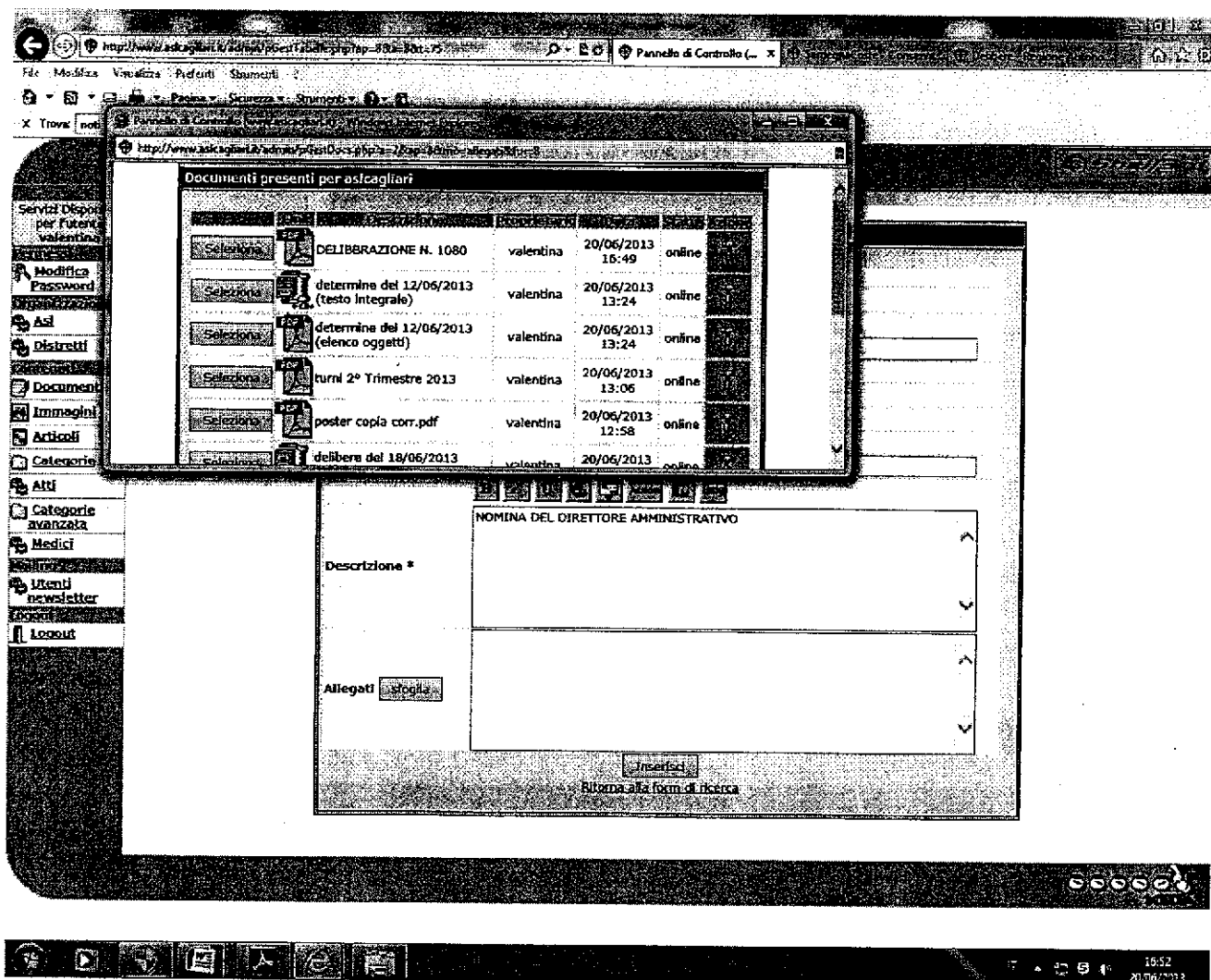
[Inserisci altri nuovi documenti](#)

Descrizione \*

NOMINA DEL DIRETTORE AMMINISTRATIVO

Allegati

[Ritorna alla form di ricerca](#)



In conclusione, le attività descritte nel presente allegato tecnico sono finalizzate a garantire al cliente la pubblicazione distinta di ogni singolo atto e degli eventuali allegati, con l'indicazione estesa del relativo oggetto e dei relativi estremi di pubblicazione.

Risultati analoghi all'output atteso dal cliente, peraltro, sono già visualizzabili sui siti internet di altre Aziende del Servizio Sanitario Regionale, che condividono la medesima piattaforma web di *back-office* per la pubblicazione degli atti.

A tal fine, pertanto, si ritiene utile riportare una schermata esemplificativa dal sito internet della ASL di Sassari.

*[Handwritten signature]*

*[Handwritten signature]*

Asl Sassari - Albo pretorio - Atti - Delibere - Archivio delibere - Mozilla Firefox

www.aslssassari.it/index.php?id=17&c=13&m=327&an=10&codex=1&id=2

aslssassari > albo pretorio > atti > delibere > archivio delibere

**ALBO PRETORIO**

- Atti
  - Delibere
    - Delibere in pubblicazione
    - Archivio delibere
      - 2013
      - 2012
      - 2011
      - 2010
      - 2009
      - 2006
      - 2007
- Determine dirigenziali
- Archivio determine
- Incarichi
- Regolamenti e Normativa
- Bandi e gare
- Concorsi e selezioni
- Bilanci

**Archivio delibere**

Ordina per: pubblicazione - scadenza

**Deliberazione del Direttore Generale n. 707 del 25/09/2013 [file.pdf]**  
Progetto "Continuità Assistenziale Diuma" - Proroga  
Data seduta: 24/09/2013  
In pubblicazione dal: 25/09/2013 - al: 09/10/2013

---

**Deliberazione del Direttore Generale n. 706 del 25/09/2013 [file.pdf]**  
Atto di Delega all'Azienda Sanitaria n. 3 di Nuoro per l'aspletamento della gara in unione di acquisto per la fornitura di energia elettrica nell'anno 2014, con opzione di rinnovo per un anno.  
Data seduta: 24/09/2013  
In pubblicazione dal: 25/09/2013 - al: 09/10/2013  
**Allegati: Atto di Delega**

---

**Deliberazione del Direttore Generale n. 705 del 25/09/2013 [file.pdf]**  
Fornitura di materiale per laparoscopia occorrente alla Farmacia Ospedaliera di Sassari, Alghero, Ozieri, Servizio Farmaceutico Territoriale e Azienda Ospedaliero Universitaria: proroga del 1/7/2013 al 31/12/2013 dei rapporti negoziali già affidati con provvedimenti nn: 808 del 27.10.2006, 879 del 15.11.2006, 592 del 25.6.2008 (successivamente rinnovati a prorogati fino al 30.6.2013), in attesa che si definisca la nuova gara a valenza regionale affidata all'ASL 3 di Nuoro con DGR 39/40 del 26.9.2012. CIG 3879223416  
Data seduta: 24/09/2013  
In pubblicazione dal: 25/09/2013 - al: 09/10/2013

---

**Deliberazione del Direttore Generale n. 704 del 25/09/2013 [file.pdf]**  
Adozione del "Regolamento preposti e referenti per il sistema di radioprotezione aziendale".  
Data seduta: 24/09/2013  
In pubblicazione dal: 25/09/2013 - al: 09/10/2013  
**Allegati: regolamento**

---

**Deliberazione del Direttore Generale n. 703 del 25/09/2013 [file.pdf]**  
"Segesta Gastoni Srl - R.S.A. San Nicola" - stipula del contratto per l'acquisizione di prestazioni di assistenza residenziale territoriale per l'anno 2013.  
Data seduta: 24/09/2013  
In pubblicazione dal: 25/09/2013 - al: 09/10/2013  
**Allegati: contratto**

---

**Deliberazione del Direttore Generale n. 702 del 25/09/2013 [file.pdf]**  
"Associazione Primavera" - stipula del contratto per l'acquisizione di prestazioni di assistenza socio sanitaria e soggetti con problematiche di dipendenza da sostanze illecite e letite per gli anni 2013/2014/2015.  
Data seduta: 24/09/2013  
In pubblicazione dal: 25/09/2013 - al: 09/10/2013  
**Allegati: contratto**

Alcune informazioni vengono inviate automaticamente a Mozilla Firefox per migliorare l'utilizzo.

start | Asl Sassari - Albo pre... | Microsoft Office... | GARA 2012 - 3000000... | Microsoft Excel - Cont... | Spool cosa condividi...

*[Handwritten signature]*

*[Handwritten initials]*