**NOME PROGETTO**

SOTTOTITOLO

# Obiettivo e ambito del documento

## Scopo del presente documento è definire le specifiche tecniche del progetto per il sistema “Gestione Gare Pubbliche”, un software web-based che consente la fruizione dei dati relativi alle gare pubbliche esposti sulla Gazzetta Ufficiale italiana ed europea.

# Premessa

## L'obiettivo del progetto è di raccogliere e memorizzare i **bandi delle gare pubbliche** sulla **Gazzetta Ufficiale**, per poi notificare gli utenti registrati al sistema della presenza di nuovi bandi, appena questi sono disponibili; gli utenti possono esprimere delle preferenze sui tipi di bando a cui sono interessati e sono notificati soltanto se i bandi rispecchiano le preferenze.

Le **Entità di Business** coinvolte dal progetto sono le pubblicazioni della Gazzetta, i bandi pubblici contenuti e il sito web della Gazzetta Ufficiale, nel quale reperire le informazioni riguardanti i bandi.

La Gazzetta è un documento ufficiale pubblicato dalla Republica Italiana con cadenza al massimo giornaliera e viene esposta in una pagina web al seguente indirizzo: <http://www.gazzettaufficiale.it/> (**Home Page della Gazzetta**).Le sue pubblicazioni sono identificate da data e un numero, generalmente fra 1 e 153, che si azzera con la fine dell'anno. Una pubblicazione espone, oltre a decreti, sentenze e regolamenti, diverse decine di bandi di vario tipo, offerti da enti anche molto diversi quali Regioni, Province, Comuni, Ministeri...

Le pubblicazioni dei bandi pubblici della Gazzetta degli ultimi 30 giorni sono esposte nel sito web della Gazzetta Ufficiale, al seguente URL: "<http://www.gazzettaufficiale.it/30giorni/contratti>" (**Pagina delle Pubblicazioni di Bandi pubblici)**.

E' possibile raggiungere la pagina delle pubblicazioni dei bandi pubblici nella sezione "Gazzetta Ufficiale" esposta nella home page, in particolare scegliendo "5a Serie Speciale - Contratti Pubblici".

Home Page della Gazzetta Ufficiale

Pagina delle Pubblicazioni di Bandi pubblici

## Gazzetta_Pub.pngNella pagina delle pubblicazioni dei bandi è possibile scegliere una pubblicazione e consultarne la relativa pagina, che contiene l'elenco dei bandi e i link agli stessi. Da questa schermata si può inoltre accedere all'archivio completo di tutte le pubblicazioni.

Dalla pagina di una pubblicazione si possono consultare i diversi bandi esposti (**Pagina di una particolare Pubblicazione**). Le informazioni che riguardano il nome del richiedente, la data di scadenza e il codice identificativo di un bando sono già disponibili in questa pagina.

Pagina di una particolare Pubblicazione

## Gazzetta_Una_Pub.pngScegliendo un particolare bando all'interno della pagina della pubblicazione, si va alla pagina del bando, dove è esposto tutto il testo del bando.

Pagina di un particolare Bando

## Gazzetta_Un_Bando.pngI bandi rappresentano delle gare pubbliche, cioè appalti di lavori o incarichi pubblici offerti dallo Stato e liberamente fruibili da privati o associazioni, previo concorso. Le informazioni principali che riguardano i bandi sono CIG (Codice Identificativo Gara) e l'oggetto della gara.

## Analizzando i bandi si nota che purtroppo non rispettano una struttura uniforme e definita. Si possono però classificare in due o tre forme, a seconda di come sono stati redatti:

* La prima forma è la meglio strutturata: c'è un titolo (che può contenere anche il CIG) e il documento è suddiviso in sezioni e sottosezioni quasi sempre uguali. In questo caso risulta facile trovare l'oggetto del bando, il suo codice e le altre informazioni utili.

## Esempio: SEZIONE II:

II.1) Descrizione.

II.1.1) Denominazione

II.1.4) Breve descrizione / oggetto dell'appalto  
II.1.5) Breve descrizione dell'appalto o degli acquisti:

SEZIONE VI: altre informazioni  
VI.3) Informazioni complementari: Codice identificativo della gara (CIG): ...

## Una seconda forma strutturata, ma suddivisa soltanto in sezioni, in cui può o meno comparire il titolo della sezione. Anche in questa serie di documenti risulta comunque non troppo complesso ricercare l'oggetto dell'appalto e il CIG (se è presente!)

Esempio:

SEZIONE II: OGGETTO DELL'APPALTO:... (CIG 7093944BC5)

SEZIONE II: Oggetto

SEZIONE II:

## L'ultima forma raccoglie tutti i documenti non strutturati e quindi quelli che presentano una descrizione unica, non divisa in sezioni. Ovviamente la difficoltà nel ritrovare le informazioni è maggiore e spesso i documenti di questo tipo non presentano neanche il CIG.

## Una volta analizzati questi documenti ne si estrae il CIG e l'oggetto del bando, per poi memorizzarli in modo persistente.

## Il sistema prevede un meccanismo di registrazione dei nuovi **utenti** con il loro username, password e la loro area privata. Ogni utente imposta alcuni parametri che rappresentano le sue **preferenze** riguardo a certi tipi di bando.

Gli utenti possono definire quali sono gli "argomenti" che interessano, cioè delle espressioni da ricercare nel testo del bando (espressioni di tipo PLUS), e quali invece sono gli argomenti trattati dai bandi di cui non si vuole essere notificati (espressioni di tipo MINUS), in modo che, se nel bando compaiono determinate espressioni, l'utente è subito scartato dalla lista dei potenziali interessati.

Quando dei nuovi bandi sono esposti sulla Gazzetta sono scaricati e memorizzati; il sistema sceglie quali sono gli utenti registrati interessati ai nuovi bandi: analizza il testo del bando e cerca se sono presenti o no le "parole chiave" definite dagli utenti, poi li notifica.

## Le **notifiche** avvengono tramite servizi Internet: posta elettronica o un sito web, pertanto in fase di registrazione l'utente definisce anche le modalità preferite di ricezione delle notifiche.

# Stati

Gli stati riguardano le entità Pubblicazione e Bando e sono necessari per la gestione interna dei dati e dei processi che coinvolgono le due entità.

Gli stati sono modificati nelle diverse fasi dei processi ed esprimono un'informazione utile alle altre fasi, che si comportano in modo diverso a seconda del valore dello stato.

Rappresentano quindi il modo in cui le fasi comunicano fra loro.

# PUBBLICAZIONE

|  |  |  |
| --- | --- | --- |
| Nome Stato | Valore | Commenti |
| STATO\_PUBBLICAZIONE | "DA\_SCARICARE" | Pubblicazione pronta da scaricare |
| "IN\_CORSO" | La pubblicazione è in fase di scaricamento / processamento |
| "SCARICATA" | La pubblicazione è stata scaricata e i suoi bandi pronti ad essere parsificati |
| "CHIUSA" | Si è concluso il processamento della pubblicazione e tutti i bandi che conteneva sono stati ricavati e memorizzati |

## Nella prima fase del processo, appena dopo il rilevamento una nuova pubblicazione e la creazione della nuova entry nella tabella pubblicazioni, il valore dello stato di questa pubblicazione è "**DA\_SCARICARE**", a significare che è pronta per essere scaricata.

Nella fase di scaricamento si prendono in considerazione soltanto le pubblicazioni il cui stato è "**DA\_SCARICARE**" (dovrebbe essere una sola). Si inizia impostando lo stato della pubblicazione a "**IN\_CORSO**", poi si scarica e si processa per ricavare i bandi che contiene. Alla fine di questa fase, quando tutti i bandi sono stati ricavati e la fase è conclusa, il valore dello stato diventa "**SCARICATA**".

La fase di controllo della pubblicazione, appena successiva alla fase di scaricamento, prende in considerazione soltanto l'ultima pubblicazione inserita, cioè quella con data più recente, e il cui stato è "**SCARICATA**". Ricerca anche se è rimasta qualche pubblicazione "**DA\_SCARICARE**": significa che una pubblicazione era pronta ma non è stata scaricata.

Successivamente, quando finisce la fase di parsificazione dei bandi, lo stato della pubblicazione è impostato a "**CHIUSA**", a significare che tutti i bandi della pubblicazione sono stati processati e che quindi non è più necessario operare in nessun modo sulla pubblicazione.

# BANDO

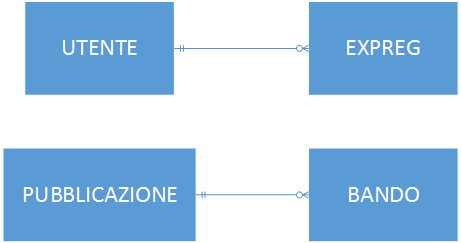
|  |  |  |
| --- | --- | --- |
| Nome Stato | Valore | Commenti |
| STATO\_BANDO | "DA\_PARSIFICARE" | Il bando è stato scaricato tutto ed è pronto per essere parsificato |
| "PARSIFICATO" | Si è conclusa l'analisi del bando ed è pronto per essere pubblicato |

## Nella fase di scaricamento della pubblicazione e in quella di controllo di completamento della pubblicazione, quando i bandi sono ricavati e scaricati dalla pubblicazione, si crea una nuova entry della tabella Bandi, con stato "**DA\_PARSIFICARE**".

### Il Parser tratta soltanto i bandi il cui stato è " **DA\_PARSIFICARE** " e procede con la loro analisi, al termine della quale lo stato del bando diventa "**PARSIFICATO**".

### Tutti i bandi " **PARSIFICATO** " sono pronti per essere pubblicati: l'ultima fase prende i bandi con stato " **PARSIFICATO** " e li analizza per decidere a quali utenti notificare i nuovi bandi.

**Modello Entità-Relazione**



**PUBBLICAZIONE**

Permette di modellare le pubblicazioni della Gazzetta Ufficiale

|  |  |  |  |
| --- | --- | --- | --- |
| Nome Proprietà | Tipo | Null / Extra | Commenti |
| CD\_PUBBLICAZIONE | BIGINT | PRIMARY, AUTO\_INCREMENT | Identificativo della pubblicazione. Popolato automaticamente |
| DATA | DATETIME |  | Data dell'inserimento della pubblicazione nel database |
| NUMERO\_PUBBLICAZIONE | SMALLINT (3) |  | Numero della pubblicazione |
| STATO\_PUBBLICAZIONE | ENUM |  | "DA\_SCARICARE", "IN\_CORSO", "SCARICATA", "CHIUSA" |
| URL | VARCHAR |  | Indirizzo URL della pubblicazione |

**Tabella 1 - Entità PUBBLICAZIONE**

# BANDO

L'entità bando rappresenta un bando di gara pubblica contenuto in una pubblicazione della Gazetta.

|  |  |  |  |
| --- | --- | --- | --- |
| Nome Proprietà | Tipo | Null / Extra | Commenti |
| CD\_BANDO | BIGINT | PRIMARY, AUTO\_INCREMENT | Identificativo del bando, generato in automatico |
| CODICE\_ESTERNO | VARCHAR (250) | UNIQUE | Codice del bando che si trova nella pubblicazione (TX17BFC9332) |
| CD\_PUBBLICAZIONE | BIGINT | REFERENCES PUBBLICAZIONE (CD\_PUBBLICAZIONE) | Rappresenta la relazione fra BANDO e PUBBLICAZIONE |
| CIG | VARCHAR (250) | Null | Codice Identificativo Gara, riportato sul bando |
| TIPO\_BANDO | VARCHAR (50) | Null | Bando di Gara / Estratto di bando di gara / Bando di gara a procedura aperta |
| TIPO\_RICHIEDENTE | VARCHAR (100) | Null | Ministeri / Regioni / Province / ... |
| NOME\_RICHIEDENTE | VARCHAR | Null | Il nome dell'ente richiedente (Associazione o privato) |
| SCADENZA | DATE | Null | La data di scadenza del bando, riportata sullo stesso |
| OGGETTO | VARCHAR | Null | Oggetto del bando |
| TESTO | TEXT | Null | Tutto il testo integrale del bando |
| URL | VARCHAR |  | Indirizzo web identificativo del bando |
| STATO\_BANDO | ENUM |  | "DA\_PARSIFICARE", "IN\_CORSO", "PARSIFICATO", "CHIUSO" |

**Tabella 2 - Entità BANDO**

Come identificativo del bando si può sfruttare il codice già assegnato dalla Gazzetta ai vari bandi (Esempio: TX17BFC9367) che si suppone essere univoco.

**EXPREG**

L'entità gestisce le espressioni regolari associate a ciascun utente, che rappresentano i parametri con cui l'utente esprime il suo interesse in determinati tipi di bando.

|  |  |  |  |
| --- | --- | --- | --- |
| Nome Proprietà | Tipo | Null / Extra | Commenti |
| CD\_EXPREG | BIGINT | PRIMARY, AUTO\_INCREMENT | Identificativo dell'espressione, assegnato in automatico. |
| CD\_UTENTE | NUMBER | REFERENCES UTENTE (CD\_UTENTE) | Rappresenta la relazione fra EXPREG e UTENTE |
| ESPRESSIONE\_PLUS | VARCHAR |  | L'espressione regolare vera e propria, di tipo PLUS |
| ESPRESSIONE\_MINUS | VARCHAR |  | L'espressione regolare vera e propria, di tipo MINUS |

**Tabella 3 - EXPREG**

Le Espressioni Regolari esprimono il fatto che l'utente sia interessato o meno in un certo argomento: le espressioni di TIPO PLUS indicano che, se nel bando compare una porzione di testo che fa match con questa, bisogna notificare l'utente. Quelle di tipo MINUS indicano invece che NON bisogna notificare l'utente e si trova qualcosa che fa match. La combinazione dei due tipi costituisce l'"argomento" di interesse.

**UTENTE**

L'entità utente rappresenta gli account che sono iscritti al sistema e che desierano ricevere delle notifiche in seguito alla pubblicazione di alcuni tipi di bando.

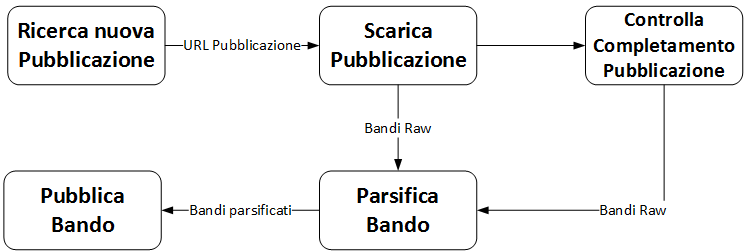
|  |  |  |  |
| --- | --- | --- | --- |
| Nome Proprietà | Tipo | Null / Extra | Commenti |
| CD\_UTENTE | BIGINT | PRIMARY, AUTO\_INCREMENT | Identificativo dell'utente, assegnato in automatico. |
| USERNAME | VARCHAR (50) | UNIQUE | Pseudonimo (univoco) dell'utente, scelto dallo stesso in fase di registrazione |
| PASSWORD | VARCHAR (50) |  | Password decisa dall'utente in fase di registrazione |
| EMAIL | VARCHAR |  | Email decisa in registrazione |

**Tabella 4 - Entità UTENTE**

Necessaria una fase di registrazione utente in cui sono inseriti tutti i dati necessari per le proprietà.

# Processi

Le fasi del processo necessarie per raggiungere l'obiettivo richiesto sono almeno 5. e varie fasi comunicano fra loro tramite dei messaggi e/o impostando uno stato, cioè proprietà delle entità Pubblicazione e Bando; unsa fase si "attiva" quando quella precedente termina.



# Ricerca Nuova Pubblicazione:

Una volta al giorno si collega al sito web della Gazzetta Ufficiale e cerca se è presente una nuova pubblicazione, controllandone la data e confrontandola con quella dell'ultima pubblicazione memorizzata dal sistema.

In caso ne sia presente una nuova, crea una nuova entry nella tabella Pubblicazione e memorizza le prime informazioni già disponibili (numero, data e URL) . Il processo può così avanzare; nel caso non sia presente nessuna nuova pubblicazione, non sono attivate le fasi successive del processo, che aspettano fino a quando la ricerca non dà un risultato positivo.

# Scarica Pubblicazione:

Questa fase del processo può essere raggiunta soltanto quando la ricerca di una nuova pubblicazione ha dato esito positivo. Essendo disponibile una nuova pubblicazione, scarica la pagina corrispondente e analizza la pubblicazione, ricavando tutti i bandi e inserendoli nel database insieme a URL corrispondente, nome dell'ente richiedente e testo per intero. Queste informazionisono già disponbili all'interno della pubblicazione.

Una volta disponibili tutti i link ai bandi, sono scaricati uno per uno: è rimossa la precedente entry corrispondente al bando e sostituita con il nuovo bando, di cui ora si conosce il testo per intero (oltre alle infirmazioni precedenti già ricavate).

Questo meccanismo permette di avere un controllo sull'effetivo scaricamento di tutti i bandi della pubblicazione.

# Parsifica Bando

Una volta disponibile il testo del bando, si parsifica e analizza per cercare di ricavare più informazioni possibili: oggetto del bando, CIG e descrizione; si separano anche logicamente nella struttura di memorizzazione. Queste informazioni, se presenti e se ricavabili, sono memorizzate all'interno dell'entità Bando, pronte per la prossima fase del processo.

Se il Parser non è stato in grado di ricavare le informazioni dal bando, perchè non presenti o male espresse o perchè il documento non era strutturato in un modo riconoscibile, rimane il testo integrale che può comunque essere utilizzato poi.

# Notifica Bando

Analizza l'oggetto del bando e eventualmente anche il testo e controlla se rispetta o no le preferenze espresse dagli utenti. Ad ogni utente sono associate diverse espressioni che devono o non devono comparire nel documento: esegue una ricerca per determinare se il documento fa match con le espressioni.

Se nessun bando interessa agli utenti registrati (secondo i criteri appena descritti), questa fase si conclude; se invece è presente almeno un bando che rispetta i parametri di uno o più utenti, li notifica. A seconda che l'utente abbia scelto di ricevere le notifiche via mail o tramite il sito web dell'applicazione, informa l'utente secondo la modalità da lui scelta.

Terminata questa fase, la catena di processi è conclusa e si deve aspettare nuovamente la disponibilità di una nuova pubblicazione prima che tutto riparta.

# Dettagli Realizzazione Tecnica

Per la realizzazione del progetto software si prevede di strutturare due differenti applicazioni:

la prima si occupa di raccogliere i dati relativi ai bandi e di memorizzarli; la seconda svolge il ruolo di analisi dei bandi e di notifica degli utenti interessati, gestendo anche tutto quello che riguarda gli utenti e la loro interazione col sistema.

L'applicazione di **RaccoltaGarePubbliche** è composta da 5 moduli che interagiscono tra loro: corrispondono circa con le fasi del processo sopra descritte.

Oltre a questi componenti, che realizzano la prima parte del sistema, va considerata l'applicazione **NotificaGarePubbliche**: questa applicazione web svolge tre ruoli principali:

* Pubblica i risultati appena elaborati: notifica un utente che è disponibile un nuovo bando (tramite sito web o posta, a seconda della preferenza dell'utente), ma solo se questo rispetta dei parametri di interesse definiti dall'utente stesso.
* Gestisce l'interfaccia che permette agli utenti di entrare nella loro area privata e di visualizzare i bandi di loro interesse.
* Permette ai nuovi utenti di registrarsi, di impostare e modificare i parametri di preferenza dei bandi e verifica gli accessi da parte di quelli registrati.