

# *Projecte “Aplicació de Resums de Premsa”*

## **Annex 1**

### **Especificació de Requeriments del Sistema**

Seu del Parlament de les Illes Balears, 22 d’octubre del 2002.

#### **Índex**

<b>ÍNDIX</b>	<b>2</b>
<b>1. INTRODUCCIÓ.</b>	<b>3</b>
<b>2. OBJECTIUS DEL SISTEMA.</b>	<b>4</b>

<b>3. CATÀLEG DE REQUERIMENTS DEL SISTEMA.</b>	<b>5</b>
<b>3.1. REQUERIMENTS D'EMMAGATZAMAMENT D'INFORMACIÓ.</b>	<b>5</b>
<b>3.2. REQUERIMENTS FUNCIONALS.</b>	<b>6</b>
3.2.1 DESCRIPCIÓ DELS REQUERIMENTS FUNCIONALS	6
3.2.2. DEFINICIÓ D'ACTORS.	11
<b>3.3. REQUERIMENTS NO FUNCIONALS.</b>	<b>12</b>
<b>GLOSSARI DE TERMES.</b>	<b>14</b>

# **1. Introducció.**

Aquest document té com a objectiu recollir els requeriments que ha de satisfer una aplicació informàtica orientada a facilitar, agilitzar i complementar l'actual procés de confecció de resums de premsa que es duu a terme a l'Àrea de Comunicació i Relacions Externes del Parlament de les Illes Balears.

## **2. Objectius del sistema.**

El desenvolupament de la nova aplicació respon a la necessitat d'assolir els següents objectius:

1. Optimitzar el procés d'elaboració de resums de premsa, automatitzant al màxim les operacions involucrades en aquesta tasca. Aquesta optimització es durà a terme tenint especial cura de minimitzar el temps emprat diàriament en aquesta operació.
2. Incorporar els processos de reconeixement òptic de caràcters (OCR) per tal de poder alimentar la base de dades amb el text íntegre de les notícies, i poder dur a terme així consultes de tipus documental sobre aquest.
3. El mòdul de consulta i recerca documental serà una aplicació web que correrà sobre la Intranet del Parlament.
4. L'aplicació permetrà la confecció de dossiers temàtics, a part dels resums diaris habituals. Els operadors podran confeccionar dossiers temàtics públics, els quals quedaran a disposició de tots els altres usuaris a la Intranet del Parlament.
5. L'aplicació contemplarà la captura de notícies publicades en diaris digitals, permetent efectuar sobre elles les mateixes operacions que es podran realitzar sobre els diaris tradicionals.
6. La publicació del resum de premsa sobre la Intranet es durà a terme com a una unitat d'informació completa, i no notícia a notícia.
7. L'aplicació donarà facilitats de visualització de les notícies en el mòdul de consulta. Ha de permetre operacions com zoom-in i zoom-out.
8. L'aplicació ha de tenir especial cura dels acabats en les operacions d'impressió de notícies. Ha de permetre operacions del tipus ajustar les imatges a les dimensions de la pàgina i incorporar encapçalaments i peus de pàgina.

## **3. Catàleg de requeriments del sistema.**

### **3.1. Requeriments d'emmagatzamament d'informació.**

En referència a la informació que s'haurà d'emmagatzemar sobre el sistema informàtic, l'aplicació haurà de satisfer els següents requeriments:

#### **1. Base de dades de notícies.**

El sistema haurà d'emmagatzemar informació referent a les notícies de premsa que formaran part del resum diari. Aquesta informació s'emmagatzemarà en una base de dades.

Concretament, les dades que s'hauran de guardar són les següents:

- Titular de la notícia
- Subtítol
- Nom de la publicació
- Data de la publicació
- Pàgina on està publicada la notícia
- Text íntegre
- Imatges associades

Aquesta base de dades serà del tipus històrica, ja que les notícies no es donaran de baixa mai de la base de dades. Per tant, la base de dades contindrà totes les notícies aparegudes al resum de premsa des del moment de la seva posada en explotació.

#### **2. Contenedor d'imatges**

El sistema emmagatzemarà també les diferents imatges escanejades associades a les notícies de la base de dades. Cada imatge estarà disponible en un arxiu d'imatge individual. Es valorarà com a opcional la possibilitat de que l'aplicació permeti operar amb imatges emmagatzemades en una base de dades multimedia.

Degut a la quantitat d'espai que requereixen les imatges, es valorarà també les possibilitats de l'aplicació quant a permetre ubicar-les sobre diferents dispositius físics: discs, cintes, DVD, etc., per tal de garantir una certa flexibilitat en aquest sentit.

### 3. Resums de premsa diaris

L'aplicació ha de permetre guardar els resums de premsa diaris en format PDF, com a documents individuals. Aquests resums tendran un període de vida de 30 dies, després del qual quedaran emmagatzemats sobre un suport off-line.

Cal comentar que les notícies del resums seguiran on-line sobre la base de dades. Només quedarà off-line el resum confeccionat en format PDF.

### 4. Dossiers de premsa

L'aplicació permetrà confeccionar dossiers de premsa temàtics, els quals estaran disponibles en dos formats: HTML sobre la intranet, i PDF. Aquests dossiers es mantindran indefinidament, tret que es decideixi donar-los de baixa. Els dossiers en format de documents PDF podran ser emmagatzemats off-line en cas de baixa.

## 3.2. Requeriments funcionals.

### 3.2.1 Descripció dels requeriments funcionals

Els requeriments funcionals de l'aplicació són els detallats a continuació:

#### A. MÒDUL DE NOTÍCIES

##### Alta de notícies

El procés d'alta de notícies es durà a terme dins del procés diari de confecció del resum de premsa.

Per donar d'alta una notícia seguirem les passes següents:

- Escanejar la pàgina o les pàgines de la notícia.
- L'aplicació executarà un procés d'autodetecció de les diferents àrees de la notícia. Sobre aquesta preselecció automàtica, l'operador rectificarà les seleccions que li semblin incorrectes, quant a la identificació dels diferents components de la notícia: títol, subtítol, text, imatges, ...
- L'operador seleccionarà les parts de la imatge que aniran al retall de premsa de la notícia en qüestió.

- L'operador informarà les dades estructurades de la notícia, o bé confirmarà les dades que li apareixin automàticament (nom de la publicació, número de pàgina, data, ...).
- Finalment, l'operador llançarà el procés d'alta pròpiament dit, el qual conté l'emmagatzamament de les imatges, la captura del text via OCR, la incorporació de la informació a la base de dades (o a l'emmagatzamament intermedi, en el cas de que l'alimentació a la base de dades es faci en diferit) i el tancament de l'operació.

En aquest procés, s'ha de permetre treballar amb imatges prèviament escanejades. També es valorarà la possibilitat d'executar determinades tasques en diferit, com, per exemple, el procés d'OCR. En aquests casos, l'aplicació hauria de mantenir un control intern de l'estat de cada notícia, quant a si té o no imatges associades, si està pendent d'OCR, etc.

Aquest és el procés bàsic del sistema, el qual serà el que necessita una major automatització per tal de poder accelerar tot el procés de confecció del resum de premsa.

### Manteniment de les notícies

A part del procés d'alta de notícies, s'ha de definir un procediment de manteniment d'aquestes notícies. Aquest no serà tan crític com l'anterior, ja que es durà a terme de forma puntual i fora del procés diari de confecció del resum de premsa.

Bàsicament, el que ha de permetre és modificar les dades associades a una notícia, així com les imatges i el text capturat via OCR.

### Baixa de notícies

També cal disposar de la possibilitat de donar de baixa una determinada notícia. Aquesta operació no es farà servir habitualment, però pot ser necessària en cas d'errors (notícies duplicades) o en processos de bateria de proves.

## B. MÒDUL DE RESUMS DE PREMSA

Aquesta operació permetrà dur a terme les següents tasques:

### Crear un resum de premsa

És l'operació inicial de confecció del resum de premsa. Per una banda, l'operador informará les dades pròpies d'un determinat resum, inicialitzarà els valors per defecte de les dades que apareixeran a cada notícia (publicació, data) i iniciarà un procés iteratiu d'alta de notícies.

El fet d'inicialitzar els valors per defecte de les dades de publicació i data (i altres) està encaminat a l'agilització del procés d'alta de notícies. L'aplicació presentarà aquests valors, fent que l'operador només hagi de confirmar-los si són correctes i rectificar-los si escau.

### Modificar resum de premsa

L'aplicació ha de permetre modificar les dades generals del resum de premsa, així incorporar-hi noves notícies (procés d'alta de notícies), esborrar notícies, i modificar notícies.

### Esborrar un resum de premsa

Aquesta operació s'emprarà en comptades ocasions, ja que, com hem dit, els resums de premsa no es donaran de baixa de la base de dades. Això no obstant, sempre és útil comptar amb la possibilitat de donar de baixa resums que s'hagin pogut crear de forma errònia, o de proves.

### Tancar un resum de premsa

Una vegada acabada la introducció de notícies o la modificació de les dades d'un resum de premsa, cal consolidar aquests canvis tancant el resum. Possiblement aquesta operació la durà a terme l'aplicació de forma implícita. L'operador tan sols haurà de sortir del resum de premsa.

### Publicar un resum de premsa

La publicació d'un resum de premsa la tractarem com a una tasca especial, apart de les tasques de creació i edició. D'aquesta manera, podrem dur a terme les validacions oportunes dels resums, així com anar configurant resums pendents de publicar, i, per tant, de validar.

Degut a la naturalesa intrínseca d'aquesta operació, aquest requeriment serà de tipus opcional, però recomanable.



## C. MÒDUL DE DOSSIERS DE PREMSA

Els dossiers de premsa són agrupacions de notícies segons un determinat criteri de selecció. Aquesta agrupació es mantindrà oberta, de manera que les notícies que s'incorporin al sistema posteriorment a la confecció d'aquest dossier, i que compleixin les condicions de selecció definides a aquest dossier, apareixeran també a aquest dossier.

### Alta de dossiers

El procés de creació d'aquest dossier seguirà les passes següents:

- Creació del dossier

L'operador crearà el dossier i informará les seves dades particular: títol del dossier, descripció, etc.

- Definició de les condicions de recerca

Aquest serà un procés iteratiu de consulta sobre la base de dades de notícies, a través del qual l'operador podrà fer refinaments successius dels criteris de selecció de notícies, fins obtenir el criteri definitiu.

- Tancament del dossier

L'operador declararà el dossier com a tancat. Es valoraran les possibles opcions d'exportació del dossier a diferents formats, com PDF, XML, etc.

### Manteniment del dossiers

Per altra banda, cal definir un mòdul de manteniment dels dossiers. En aquest mòdul, l'operador ha de permetre modificar les dades generals del resum, així com variar les condicions de recerca de notícies.

### Baixa de dossiers

També cal preveure una operació de baixa de dossiers, encara que inicialment no es contempli donar-los de baixa. Com en casos anteriors, pot ser útil per poder esborrar dossiers de prova o duplicats.

## Publicació de dossiers

De la mateixa manera com ho hem definit per als resums de premsa, també distingirem l'operació de publicació de dossiers de les operacions de creació i edició de dossiers de premsa. Podrem anar creant dossiers de premsa, els quals seran publicats una vegada hagin estat oportunament validats.

Com en el cas dels resums de premsa, considerarem aquest requeriment com a opcional.

## D. MÒDUL DE RECERCA I VISUALITZACIÓ DE NOTÍCIES

Aquest mòdul és també un component crític de l'aplicació, ja que és el que ens permetrà fer ús de la informació introduïda al sistema.

La consulta de notícies estarà disponible sobre la intranet del Parlament. Per tant, aquest mòdul serà una aplicació web.

Per llançar consultes sobre la base de dades de notícies de premsa, un usuari haurà de seguir les següents operacions:

- Selecció de les condicions de recerca

Sobre una pantalla on apareixeran les diferents possibilitats d'interrogació, l'usuari definirà uns criteris de selecció de notícies. És a dir, informarà els valors que vol localitzar en determinats camps de la base de dades. Una vegada fet això, llançarà el procés d'interrogació al sistema.

- Visualització de les notícies seleccionades

Com a conseqüència del pas anterior, l'usuari obtindrà una llista de titulars de les diferents notícies de premsa que compleixen les condicions de recerca prefixades, en una ordenació predefinida.

Des d'aquest punt, l'usuari podrà generar un llistat amb les notícies de la llista, visualitzar una de les notícies de la llista, i, opcionalment, exportar aquesta selecció a un document de un determinat format (PDF, XML).

- Visualització d'una notícia

Fent un click sobre una de les notícies de la llista anterior, l'usuari podrà visualitzar tota la informació vinculada a ella: dades estructurades, imatges associades, etc.

Es valorarà l'agilitat del sistema en termes del temps de resposta, ja que aquest ha de ser acceptable per a que l'aplicació sigui operativa per a l'usuari final.

Des d'aquesta pantalla de visualització, també podrà dur a terme operacions com la impressió de la notícia.

## E. MÒDUL D'ADMINISTRACIÓ DE L'APLICACIÓ

Aquest mòdul està orientat al manteniment de les opcions generals de l'aplicació. Això fa referència a contingut de taules mestre (categories de notícies, nom de publicacions, etc.), preseleccions de variables, criteris d'ordenació, etc.

### 3.2.2. Definició d'actors.

Els diferents actors que interactuaran amb el sistema són els següents:

- Operador

S'encarregarà de les operacions de manteniment de notícies, resums i dossiers de premsa.

- Administrador

S'encarrega de dur a terme les tasques de manteniment de taules mestres i de les opcions generals de configuració de l'aplicació, contingudes en el mòdul d'administració de l'aplicació.

- Usuari de consulta

Tindrà accés al mòdul web de consulta i visualització de notícies.

### 3.3. Requeriments no funcionals.

L'aplicació haurà de satisfer els següents requeriments NO funcionals:

#### Entorn d'execució

El mòdul de gestió d'aquesta aplicació haurà de funcionar sobre dues estacions de treball simultàniament, les quals són ordinadors del tipus PC operant sobre el sistema operatiu Microsoft Windows 2000, connectats a una xarxa TCP/IP contra servidors de la família Microsoft Windows 2000 Server.

La base de dades sobre la que es recolzarà l'aplicació serà una de les actualment existents al sistema informàtic del Parlament, les quals són:

- Microsoft SQL Server 2000
- Progress 9.1
- Lotus Notes Domino

Els escaners emprats són Epson GT-12000.

El mòdul de consulta web funcionarà sobre Microsoft Internet Information Server 5 i/o Lotus Notes Domino.

#### Accessibilitat de les dades

La base de dades haurà de facilitar l'accessibilitat de la informació emmagatzemada des d'interfícies obertes, com és el cas d'OLE-DB, ADO i ODBC.

Aquest requeriment és molt recomanable des del punt de vista d'integració d'aplicacions dins el sistema. Això no obstant, s'ha definit com a requeriment opcional.

#### Arquitectura de 32-bits

L'aplicació estarà construïda internament amb codi de 32-bits, per tal de poder obtenir un rendiment òptim sobre l'arquitectura actual del sistema informàtic, que també és de 32-bits.

### Continuïtat del producte

L'empresa proveïdora de l'aplicació ha de garantir una continuïtat del producte durant un període no inferior a 4 anys (1 legislatura), o bé garantir l'existència d'una alternativa viable, en cas de descatalogació.

### Demostració de l'aplicació in-situ

Les empreses concursants duran a terme una demostració de la seva aplicació o maqueta a la Seu del Parlament de les Illes Balears.

### Borsa d'hores destinades a la personalització de l'aplicació

Les empreses concursants inclouran en la seva oferta una borsa d'hores de feina destinades a tasques de personalització i adaptació de l'aplicació a les necessitats de canvi detectades durant el primer any d'explotació de l'aplicació.

### Curs de formació per a usuaris i administradors

Les empreses concursants inclouran en la seva oferta un curs de formació d'usuaris destinat a l'operativa habitual de l'aplicació, i un altre destinat a l'administració del producte. Aquesta formació es durà a terme al Parlament de les Illes Balears.

### Període de garantia

El període de garantia ha de tenir una durada mínima d'un any, a comptar des de la finalització de la instal·lació i posada en funcionament de l'aplicació.

La garantia ha d'incloure l'adequació de l'aplicació als requeriments funcionals, d'operació i de seguretat descrits en aquest document.

## Glossari de termes.

**Base de dades documental:** Anomenam base de dades documental a una base de dades que conté un o més camps d'informació poc o gens estructurada. Aquests camps poden ser del tipus *long-text*, és a dir, camps que contenen cadenes de caràcters de gran longitud. Els registres d'aquest tipus de bases de dades esdevenen documents, degut a la gran quantitat d'informació emmagatzemada en cada un d'ells. Sobre aquest contingut, els sistemes gestor de bases de dades documentals implementen operacions especials, destinades a aquest tipus de contingut, orientades principalment a la recerca d'informació.

**Consulta documental:** Les consultes documentals són recerques d'informació dutes a terme sobre un fons documental, gestionat normalment per una base de dades documental. Es caracteritzen per ser consultes per text lliure, enlloc de consultes per identificadors o referències. Poden incloure un o més *thesaures*, els quals permeten relacionar termes de consulta des del punt de vista semàntic.

**Diari digital:** Anomenam diari digital a les publicacions de premsa que són servides sobre medis electrònics digitals, com són les pàgines web. Aquests diaris ja no fan ús del paper imprès en la seva edició, sinó que apareixen publicats als *web-sites* dels diaris tradicionals. La seva aparició ha vengut motivada pel gran avanç que ha tengut Internet en els darrers anys.

**Dossier de premsa:** Anomenam dossier de premsa a un conjunt de retalls de premsa que compleixen un determinat criteri de selecció. Aquests criteris solen ser de tipus temàtic. Recullen les notícies publicades sobre un determinat tema o fet ocorregut.

**Resum de premsa:** Un resum de premsa és un conjunt de retalls de premsa de notícies publicades durant un determinat període de temps, normalment el mateix dia. El tipus de notícies que queden incloses al resum de premsa determina el perfil d'aquest resum, i es defineix d'acord a l'interès temàtic dels usuaris als que va destinat aquest resum de premsa.

# **ANNEX-2: Taula normalitzada de característiques i funcionalitats del sistema d'informació**

Seu del Parlament de les Illes Balears, 22 d'octubre del 2002.

---

Pel Servei de Comunicació i Relacions Externes:  
Vist-i-plau

Francesca Sancho i Lladó

---

Pel Servei d'Informàtica:  
Vist-i-plau

Joan Miquel Tramullas i Mesquida

---

ÀREA	Requeriments / Característiques	Comentaris	Valoració	Tipus Valor
<b>BASE DE DADES</b>	Treballa sobre MS SQL SERVER 2000			(S/N)
	Treballa sobre PROGRESS 9.1			(S/N)
	Treballa sobre LOTUS NOTES DOMINO			(S/N)
<b>IMATGES</b>	Permet enllaçar imatges contingudes a una BD			(S/N)
	Permet enllaçar imatges ubicades a cintes, DVDs, ...			(S/N)
<b>ESCAINEIG</b>	Procés d'escaneig integrat			(S/N)
	Permet integrar imatges prèviament escanejades			(S/N)
	Autodetecció d'àrees escanejades			(S/N)
	Tipificació d'àrees escanejades (titular, subtítol, ...)			(S/N)
<b>OCR</b>	Procés d'OCR integrat			(S/N)
	Incorporació de diccionari en català per a l'OCR			(S/N)
	Incorporació de diccionari en castellà per a l'OCR			(S/N)
	Incorporació de diccionari en altres idiomes per a l'OCR			(S/N)
	Incorporació de diccionari en altres idiomes per a l'OCR			(S/N)
<b>ALIMENTACIÓ BD</b>	Alimentació a la base de dades diferida			(S/N)
	Alimentació a la base de dades on-line			(S/N)
<b>AUTOMATITZACIÓ FLEXIBILITAT D'OPERACIÓ</b>	Grau d'automatització del procés			(0-10)
	Permet efectuar tasques en diferit (imatges, OCR,...)			(S/N)
	Control de l'estat d'una notícia (pendent OCR, ...)			(S/N)
	Permet modificar dades de la notícia			(S/N)
	Permet modificar imatges de la notícia i relançar OCR			(S/N)
	Fa un seguiment de l'estat de la notícia			(S/N)
	Permet donar notícies de baixa			(S/N)



ÀREA	Requeriments / Característiques	Comentaris	Valoració	Tipus Valor
	Baixa efectiva de la notícia al moment d'abandonar l'aplicació			(S/N)
	Permet vincular una notícia a un altre resum, desvinculant-la del primer			(S/N)
	Permet esborrar un resum de premsa sense esborrar les notícies			(S/N)
	Permet la recuperació de notícies esborrades			(S/N)
<b>MULTIUSUARI</b>	Treball multiusuari sobre un mateix resum de premsa			(S/N)
	Treball multiusuari sobre un mateix dossier			(S/N)
<b>PUBLICACIÓ</b>	Existeix l'operació publicació de resum de premsa com a operació aïllada de l'edició del resum			(S/N)
	Existeix l'operació publicació de dossier com a operació aïllada de l'edició del dossier			(S/N)
<b>EXPORTACIÓ</b>	Permet exportar un resum a format PDF			(S/N)
	Grau de l'exportació del resum a PDF (indexs, links, ...)			(0-10)
	Permet exportar un dossier a format PDF			(S/N)
	Grau de l'exportació del dossier a PDF (indexs, links, ...)			(0-10)
	Permet exportar un resum a format XML			(S/N)
	Permet exportar un dossier a format XML			(0-10)
	Permet enviar els resultats d'una consulta directament per e-mail			(S/N)
<b>CONSULTA WEB</b>	Consulta web de resums publicats			(S/N)
	Consulta web de dossiers publicats			(S/N)
	Mòdul web de consulta documental de notícies			(S/N)
	Possibilitat de refinaments successius dels termes de la consulta documental			(S/N)

ÀREA	Requeriments / Característiques	Comentaris	Valoració	Tipus Valor
	Exportació a PDF dels resultats de la consulta			(S/N)
	Exportació a XML dels resultats de la consulta			(S/N)
	Versió per a imprimir dels resultats de la consulta			(S/N)
	Possibilitat d'enviar directament els resultats de la consulta per e-mail			(S/N)
<b>ADMINISTRACIÓ</b>	Existència d'un mòdul d'administració dels paràmetres de l'aplicació			(S/N)
	Criteri d'ordenació dels llistats configurable per l'usuari			(S/N)
	Existència d'un mòdul de manteniment de les publicacions de premsa			(S/N)
<b>MODULARITAT</b>	Existència de mòduls addicionals			(S/N)
	Enumeració de mòduls addicionals sol·licitats			
<b>NO-FUNCIONALS</b>	Compatibilitat amb escaners Epson GT-12000			(S/N)
	Compatibilitat amb Microsoft IIS 5.0			(S/N)
	Compatibilitat amb Lotus Notes Domino (web)			(S/N)
	Permet accés a les dades de la BD des d'OLE-DB			(S/N)
	Permet accés a les dades de la BD des d'ODBC			(S/N)
	Permet accés a les dades de la BD des d'ADO			(S/N)
	Arquitectura de 32-bits			(S/N)
	Garantia de continuïtat del producte de 4 anys			(S/N)
	Grau de viabilitat de la solució alternativa en cas de descatalogació del producte			(0-10)

ÀREA	Requeriments / Característiques	Comentaris	Valoració	Tipus Valor
	Demostració del producte al Parlament Tamany de la borsa d'hores-home incloses en l'oferta Preu de les hores-home extres de personalització de l'aplicació Hores del curs de formació d'usuaris Hores del curs de formació a administradors Duració del període de garantia			(S/N)