

# Gestor de documents

## Manual d'usuari Dockitty

Q1 2022-23 | Subgrup 31.2

Pere Carrillo

Pol Garrido

Marc Ordóñez

Laura Pérez

<b>1. Introducció</b>	<b>4</b>
<b>2. Menú Principal</b>	<b>5</b>
Nou Document	5
Carrega Document	5
Gestor Documents	6
<b>3. Modificar Document</b>	<b>7</b>
Títol, Autor, Contingut	7
Desar	7
Exportar	8
Nou Document	9
Importar Document	9
Eliminar Document	10
Esborrar Camps	10
<b>Carregar repositori</b>	<b>11</b>
<b>4. Gestor de Documents</b>	<b>12</b>
Mostrar Contingut	13
Etiquetes	13
Associar etiquetes a documents	14
Botons laterals	15
Mostra tots	15
Modificar	15
Crear	16
Exporta	16
Llistar títols d'un autor	16
Llistar autors per prefix	17
Cerca k documents semblants a un document	18
Cerca k documents rellevants donades unes paraules	19
Consultes Booleanes	20
<b>5. Gestor de Consultes Booleanes</b>	<b>21</b>
Tipus d'input	21
Cerca	22
Desar	23
Modificar	24
<b>6. Accions globals</b>	<b>25</b>
Exportar Document	25
Tornar al menú principal i tornar enrere	25
Carregar el nostre repositori	25

<b>7. Informació addicional</b>	<b>26</b>
Funcionalitats generals	26
Format propietari .kitty	26
Estructura de carpetes	27
Pesos de l'espai vectorial	28
Consultes Booleanes	30

# 1. Introducció

Dockitty és una eina programada en Java que té com a objectiu la gestió de documents, és a dir, ser capaç de crear nous documents i modificar-los; i facilitar la cerca i navegació entre ells aplicant diferents tipus de filtres.

També ens permetrà importar i exportar els documents amb tres tipus de format:

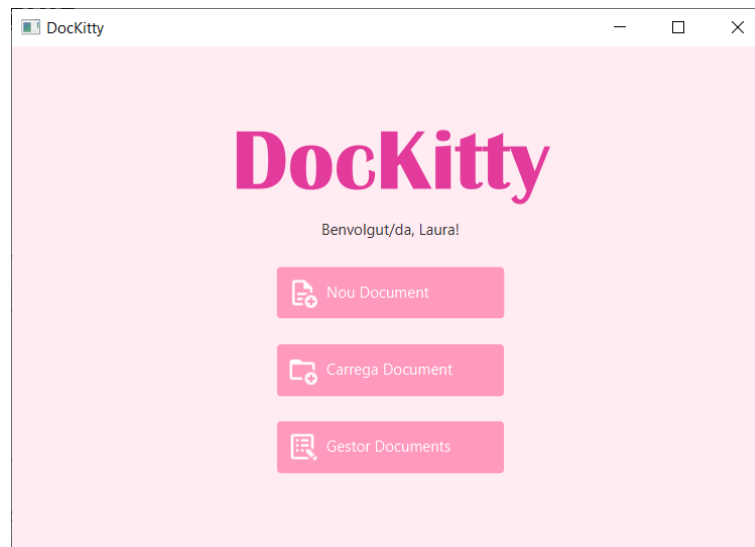
- Text pla ([.txt](#))
- Extensible Markup Language ([.xml](#))
- Format Dockitty ([.kitty](#))

El més important d'aquest gestor de documents, és que ens permet realitzar consultes sobre les obres registrades a l'aplicació:

- Ens permet obtenir el llistat d'autors donat un prefix.
- Els títols d'un autor.
- Consultes de semblança de documents i de rellevància per un conjunt de paraules.
- Consultes booleanes sobre els documents indicant les paraules o conjunts de paraules que volem que apareixen o no a les frases dels documents.

També ens permet afegir etiquetes als documents per facilitar la gestió.

## 2. Menú Principal



*Figura 1. Menú Principal*

El primer que es veu a l'hora d'executar el programa és el Menú principal [Figura 1], una vista molt senzilla, composta per el nom de l'aplicació, un text de benvinguda personalitzat i tres botons que compleixen les següents funcionalitats:

1. Crear un nou document
2. Carregar un document
3. Anar al gestor de documents

### Nou Document

Crea un document buit i es carrega la pantalla de [Modificar Document](#) amb tots els camps buits.

### Carrega Document

El segon botó serveix per carregar un o més documents de l'ordinador al nostre sistema. Es mostrarà una pantalla per seleccionar els documents que vulguis, seran llegits i carregats al sistema.

Si un document ja existia, és a dir, comparteix títol i autor amb un altre document, pots escollir si sobreescriure'l o no.

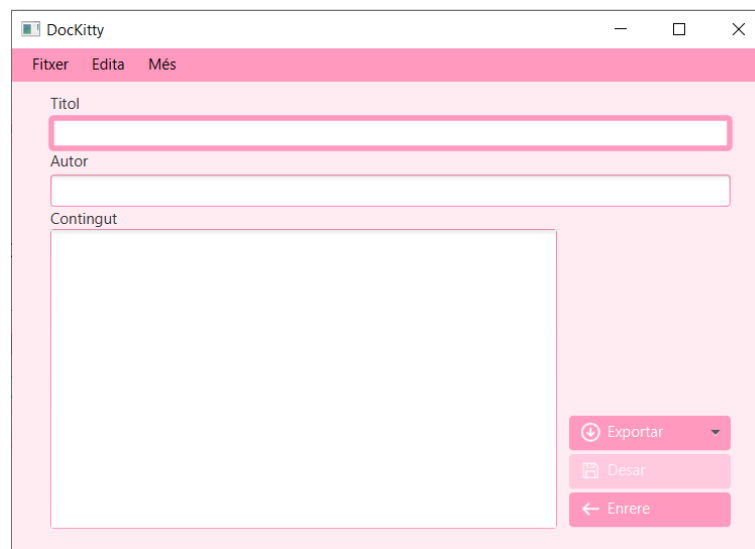
Si has escollit un sol document, podràs decidir modificar-lo o no, i se't carregarà la pantalla a [Modificar Document](#) amb les dades del document als camps corresponents.

En cas de carregar un document que no tingui títol o autor, saltarà un avís per avisar a l'usuari de l'error i no es carregarà aquell document.

## Gestor Documents

Finalment, la tercera opció és carregar el [Gestor de Documents](#), que et portarà a la pantalla on podràs fer cerques i operacions amb els documents que tinguis al sistema.

### 3. Modificar Document

The image shows a web browser window titled 'DocKitty'. The interface has a pink header bar with three menu items: 'Fitxer', 'Edita', and 'Més'. Below the header, there are three input fields: 'Títol' (Title), 'Autor' (Author), and 'Contingut' (Content). The 'Contingut' field is a large text area. To the right of the 'Contingut' field, there are three buttons: 'Exportar' (Export) with a download icon, 'Desar' (Save) with a floppy disk icon, and 'Enrere' (Back) with a left arrow icon. The entire form is set against a light pink background.

*Figura 2. Modificar Document*

Aquesta pantalla serveix per modificar documents. Tant de nous com d'existents al sistema. Com es pot veure, està format per tres grans camps on s'hi pot escriure el Títol, l'Autor i el Contingut.

#### Títol, Autor, Contingut

En aquests camps es mostrarà la informació del document i el podràs editar. Amb la informació introduïda podràs fer diferents accions.

#### Desar

La funcionalitat més important és *desar*<sup>1</sup>, que emmagatzemarà, o bé actualitzarà, el document amb el títol, l'autor i el contingut que hagi introduït al sistema. En el cas de *desar* un document al que se li ha modificat el títol i l'autor de manera que coincideix amb un altre amb el mateix identificador, se't demanarà si vols sobre escriure o no l'antic document.

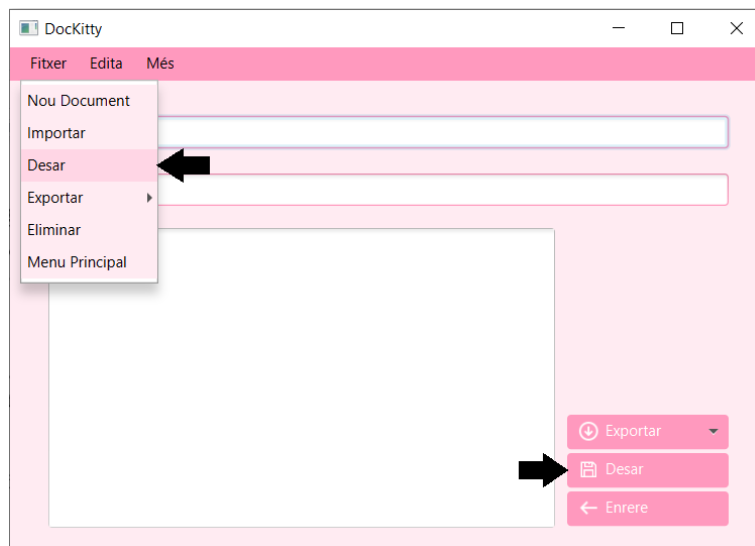
Ho pots fer amb el botó de la figura o bé a la barra de dalt Fitxer > Desar.

Si no has introduït el camp títol o autor saltarà un avís demanant que introdueixis títol i autor.

Si vas enrere o surts d'aquesta vista i tens informació sense *desar* se t'avisarà per tal que decideixis si vols guardar els darrers canvis o deixar-ho com estava.

---

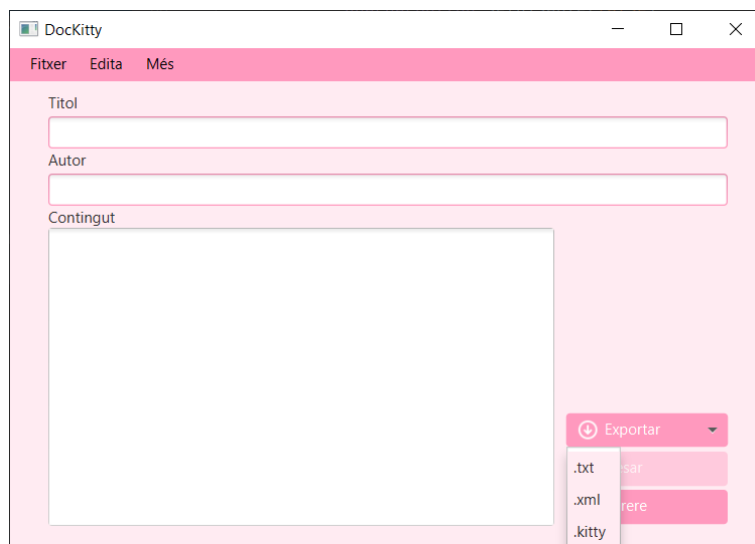
<sup>1</sup> És obligatori que tots els camps estiguin emplenats.  
La suma de caràcters del títol i l'autor ha de ser menor a 200.



*Figura 3. Opcions per desar*

## Exportar

Si vols exportar el document que estàs editant, pots fer-ho amb el botó de la figura [Figura 4] o bé a Fitxer > Exportar [Figura 5]. Se't desarà el document al programa i posteriorment farà l'acció d'[Exportar el document](#).



*Figura 4. Exportar document des del botó*



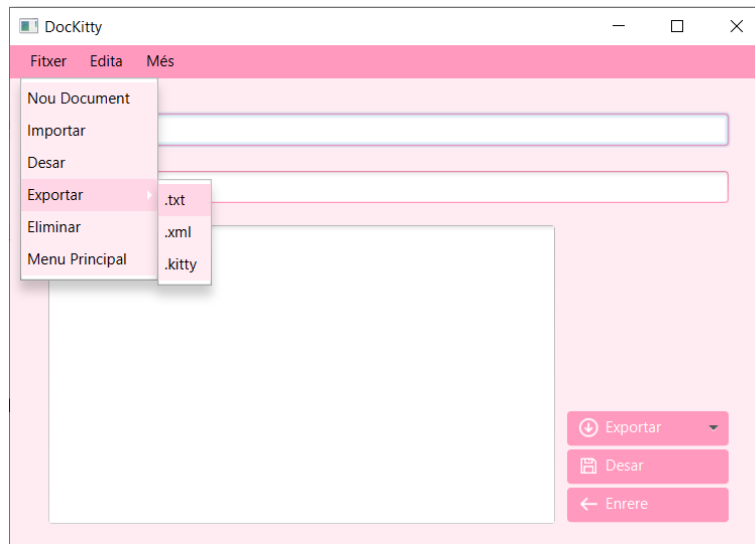


Figura 5. Exportar document des del menú

## Nou Document

Vés a la barra de dalt, al desplegable Fitxer > Nou Fitxer [Figura 6]. S'obrirà un nou fitxer amb els camps buits. Si vols crear un nou document i no has desat l'anterior se t'avisarà abans d'obrir un nou document perquè puguis desar-lo.

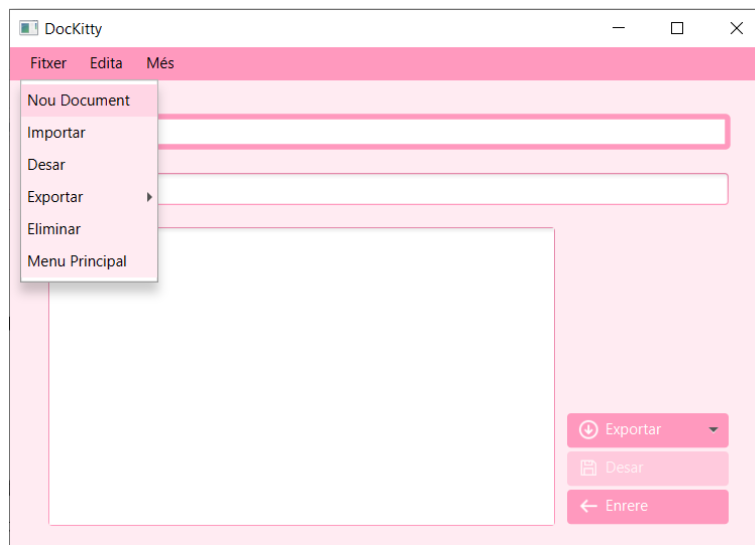


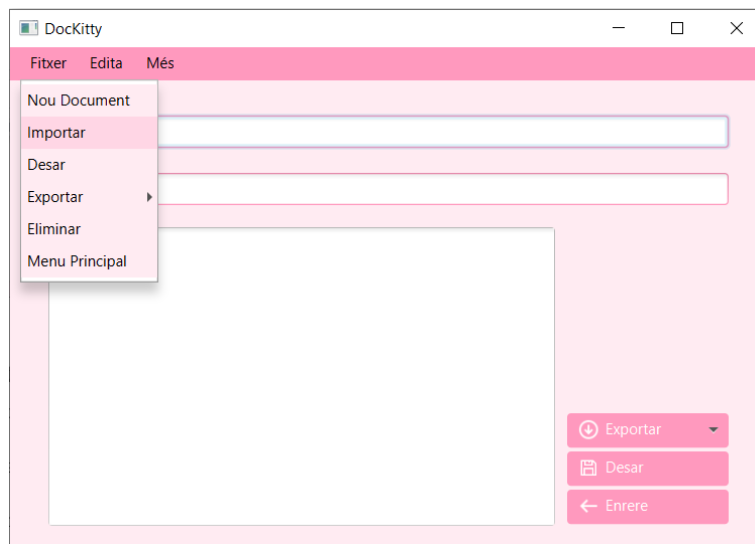
Figura 6. Crear un nou document

## Importar Document

Per importar un document, clica Fitxer > Importar. Es mostrarà una pantalla per seleccionar el document que vulguis, serà llegit, carregat al sistema i se't mostrarà el seu contingut als camps corresponents<sup>2,3</sup>.

<sup>2</sup> Si un document ja existia, és a dir, comparteix títol i autor amb un altre document, pots escollir si sobreescriure'l o no.

<sup>3</sup> Si un document no té títol o autor, no es podrà carregar i s'avisarà a l'usuari.

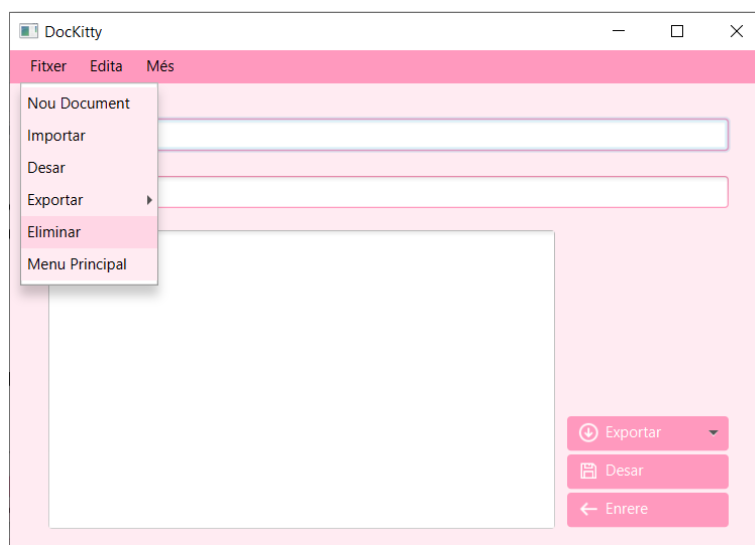


*Figura 7. Importar document des de l'editor*

Un cop s'hagi acabat la importació es notificarà a l'usuari.

## Eliminar Document

Si per algun motiu vols eliminar el document que estàs editant, vés a la barra de dalt Fitxer > Elimina Document<sup>3</sup> i confirma.



*Figura 8. Eliminar el document*

## Esborrar Camps

Si només vols eliminar el que tens escrit però mantenir el document, pots anar a Edita > Esborra tot, que et permetrà eliminar els camps que estiguis editant. Si vols sortir sense desar, et demanarà confirmació per desar o no l'estat actual del document

<sup>3</sup> Tingues en compte que s'eliminarà l'última versió desada, no la que se't mostri en pantalla

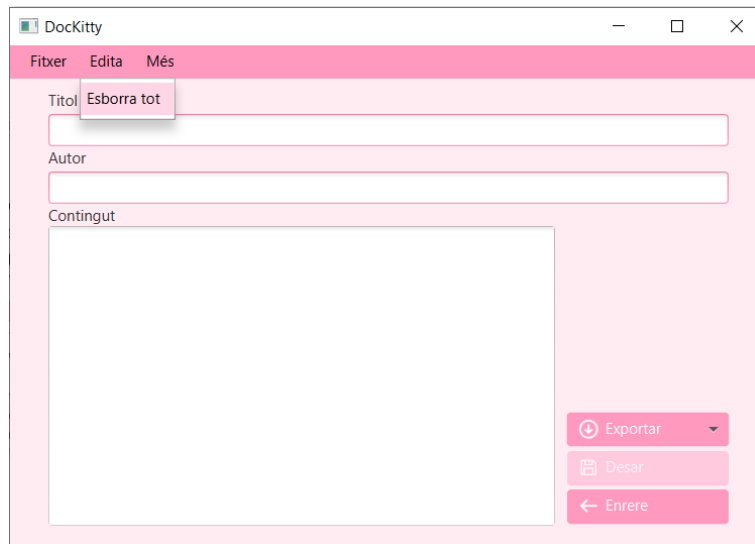


Figura 9. Esborrar els camps

## Carregar repositori

Per veure el codi font de la nostra aplicació pot accedir al repositori de GitLab des de Més > GitLab.

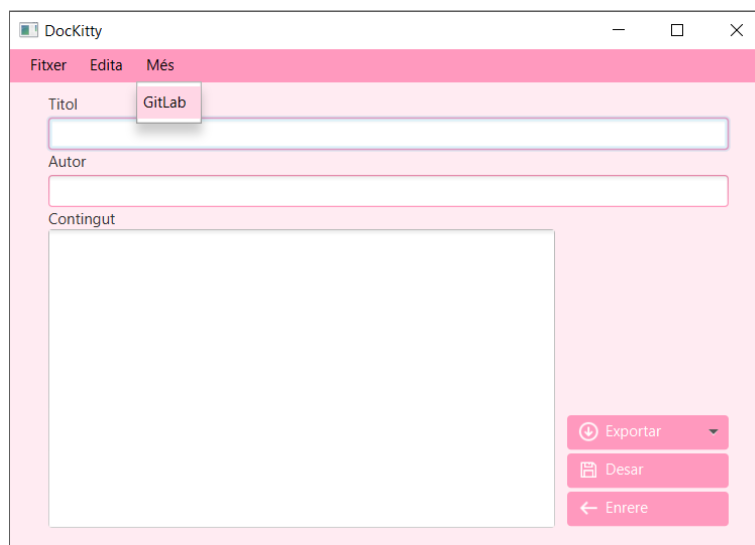


Figura 10. Carregar Gitlab

## 4. Gestor de Documents

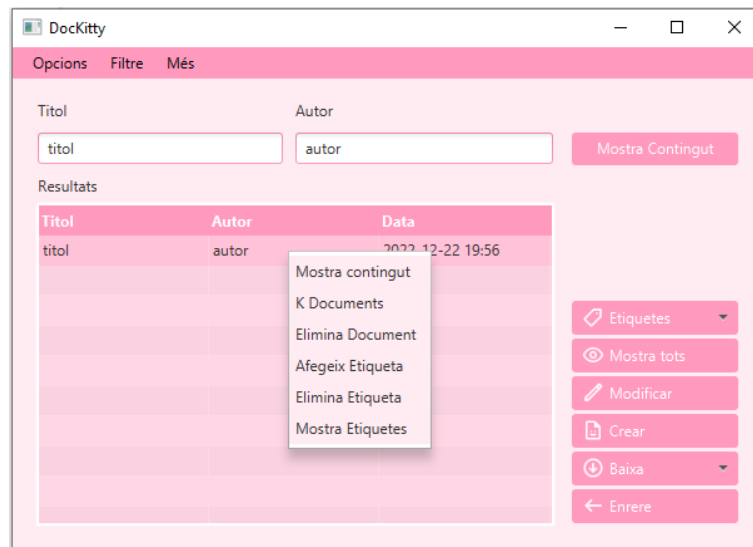


Figura 11. Gestor de documents per defecte

El gestor de documents et permet fer consultes i aplicar filtres als documents per obtenir diferents conjunts de resultats.

El primer que es veu a l'hora de carregar aquesta pàgina és la [Figura 11].

Té un camp principal on es van mostrant tots els resultats que compleixen certes restriccions. Per defecte mostra tots els documents del sistema.

Els dos camps de dalt, (Títol i Autor) serveixen per operar amb un document en concret, escrivint el seu títol i autor als camps corresponents i ens mostrarà el contingut, en cas que existís.

Depenent del filtre que apliquis, veuràs diferents camps i botons, i depenent del resultat que retornin, se't mostrarà en format de taula o de llista. La manera més directa per escollir el filtre és anar al desplegable Filtre [Figura 12] de la barra superior i seleccionar el que vulguis utilitzar.

L'opció Més, com a totes les vistes (Més > GitLab), permet accedir al repositori del projecte.

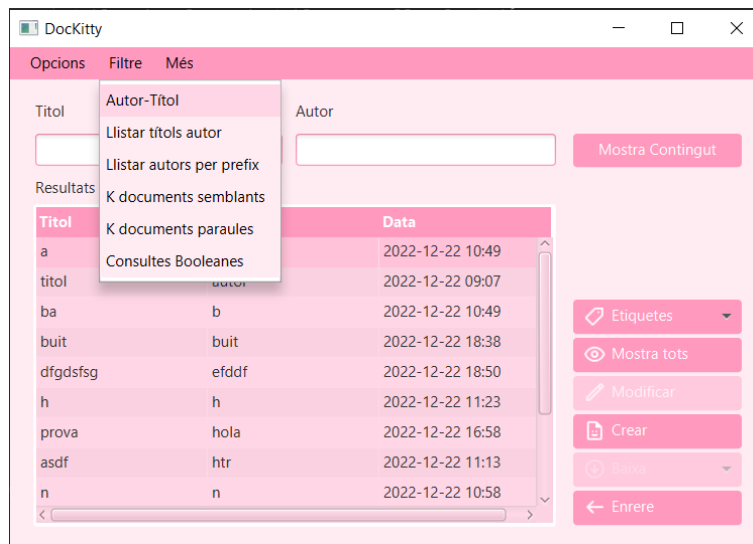


Figura 12. Menú de filtres

## Mostrar Contingut

El primer mode, i el que està per defecte quan carregues aquesta vista, serveix per mostrar el contingut d'un document, sense poder editar-lo.

Hi ha diferents formes d'utilitzar-ho:

- Emplenar els camps manualment i clicar a "Mostra Contingut"
- Clica el document que vulguis veure a la taula o a la llista i prémer el botó de "Mostra Contingut".
- Fent click dret a un document de la taula o de la llista, apareixerà l'opció de mostrar el contingut, clica i el podràs veure.

## Etiquetes

En clicar el botó "Etiquetes" es mostren les etiquetes definides al sistema, es llistaràn totes les etiquetes que hakis creat, la de "Favorits" ve per defecte i no la podràs eliminar. A més, també apareixen les opcions "Afegir Etiqueta" i "Eliminar Etiqueta".

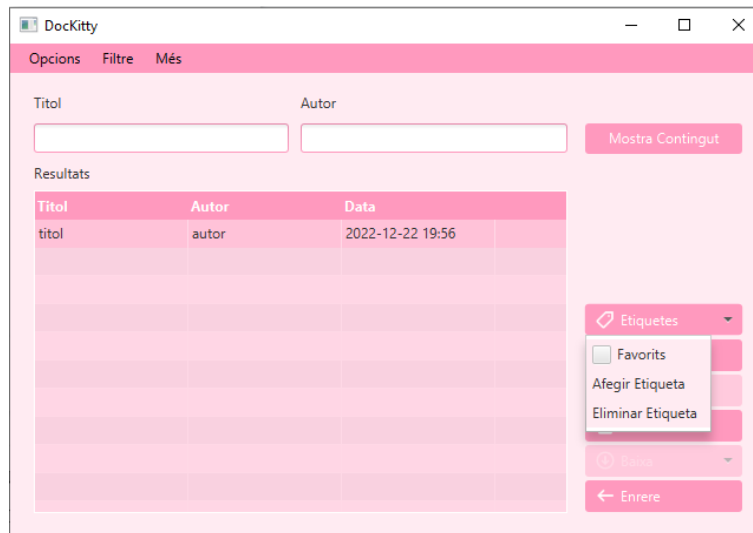


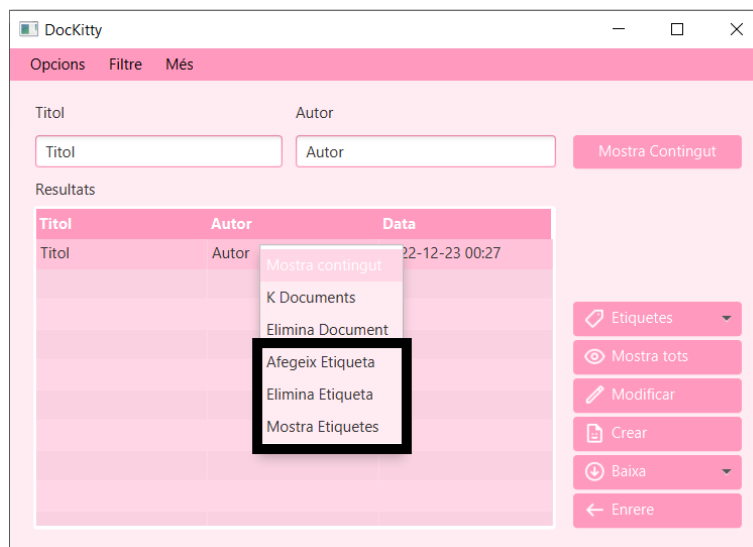
Figura 13. Opcions d'etiqueta

Si selecciones l'opció “Afegir Etiqueta”, apareixerà un *popup* per definir el nom de la nova etiqueta i les opcions d'acceptar per desar la nova etiqueta o bé cancel·lar per no desar-la. Si afegeixes una etiqueta ja existent, es mantindrà l'antiga per tal de no eliminar les obres.

Si en canvi selecciones “Eliminar Etiqueta”, se't mostrarà la llista d'etiquetes que es poden eliminar, selecciona la que vulguis i clica a acceptar per eliminar-la o cancel·lar per no fer-ho<sup>4</sup>. L'etiqueta “Favorits” no podrà ser eliminada. Directament no apareixerà en la llista d'etiquetes a eliminar.

### Associar etiquetes a documents

Per treballar amb etiquetes i documents pots fer clic dret al document que vulguis i tindràs aquestes tres opcions.



<sup>4</sup> En cas de només existir l'etiqueta de l'etiqueta “Favorits”, veuràs que no podràs eliminar cap etiqueta.

Figura 14. Afegir i eliminar etiqueta d'un document

Per associar una etiqueta a un document fes clic dret sobre el document desitjat a la taula i selecciona l'opció “Afegeix etiqueta”, apareixerà un popup amb les etiquetes que hi ha al sistema i podràs escollir quina afegir<sup>5</sup>. Únicament se't mostraran les etiquetes que encara no conté el document, per evitar afegir-ne de redundants.

Si per el contrari vols eliminar una etiqueta d'un document, selecciona “Eliminar etiqueta” i escull l'etiqueta que ja no vulguis<sup>6</sup>. Únicament se't mostraran les etiquetes que conté el document, ja que són les úniques que podràs eliminar.

Per mostrar les etiquetes associades a un document, selecciona l'opció “Mostra Etiquetes”<sup>7</sup>.

Una vegada definides i associades les etiquetes, sempre podem filtrar els documents clicant al botó “Etiquetes” i seleccionant les etiquetes per les que volem filtrar activant les *checkbox* o desactivant-les. En el moment de filtrar es mostraran aquells documents que tinguin *alguna* de les etiquetes del filtre, sense necessitat de que el document contingui totes les etiquetes alhora.

## Botons laterals

### Mostra tots

Com el seu nom indica, mostra tots els documents registrats al sistema a la taula interactiva, aquesta funcionalitat s'activa per defecte per agilitzar la gestió dels documents. En cas de voler activar-la al acabar una consulta o similar, només cal clicar el botó “Mostra Tots”. Si tens seleccionat alguna etiqueta per filtrar i vols que apareguin tots els documents de nous, pots donar-li a “Mostra Tots” o anar a “Etiquetes” i desmarcar totes les etiquetes. En cas de donar-li a “Mostra Tots”, se't desmarcaran automàticament.

### Modificar

Per modificar un document des del Gestor de Documents tenim dues opcions.

- Fer clic sobre el document a la taula i clicar el botó “Modificar”.
- Fer doble clic sobre el document a la taula.

Ambdues omplen automàticament els camps de títol, autor i contingut, permetent la modificació directa dels seus paràmetres.

Se'ns obrirà la vista [Modificar Document](#).

---

<sup>5</sup> No podràs afegir més etiquetes de les ja existents al sistema

<sup>6</sup> Si el document no té cap etiqueta associada i intentem eliminar el sistema avisa de que el document no té cap etiqueta associada

<sup>7</sup> Si el document no té cap etiqueta associada no es mostrarà cap etiqueta

## Crear

Si es prem el botó “Crear”, s’obrirà la pantalla de modificar documents amb tots els camps buits.

## Exporta

Permet descarregar en qualsevol dels tres formats suportats (.txt,.xml,.kitty) el fitxer seleccionat. Per fer-ho selecciona el document que vols exportar fent clic sobre ell a la taula i després prem el botó “Exporta”. Finalment s’ha de seleccionar el format en què es vol descarregar i posteriorment es durà a terme l’acció d’[Exportar el document](#).

## Llistar títols d’un autor

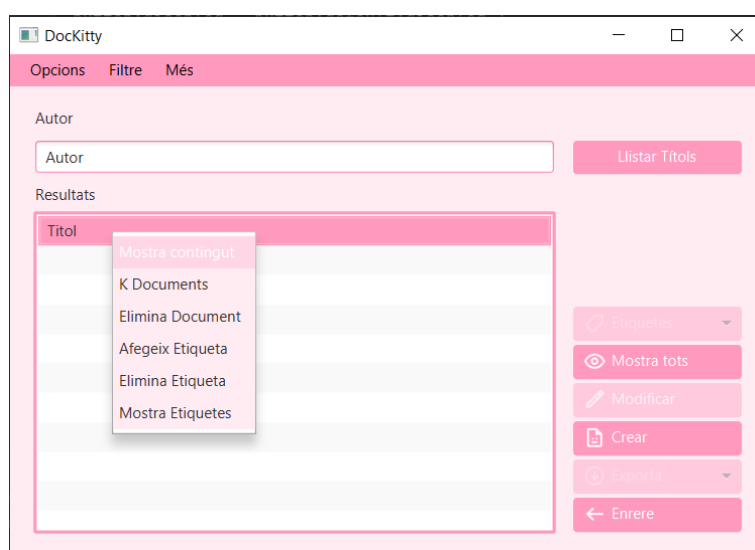


Figura 15. Resultat de llistar títols d’un autor

En aquest tipus de cerca tenim un camp per introduir el nom d’un autor. Com el seu nom indica, et permet llistar els títols dels documents del l’autor demanat<sup>8</sup>. S’hi pot accedir a través de “Filtre” i seleccionant l’opció “Llistar títols autor”.

Per utilitzar-ho, simplement introdueix el camp “Autor” i clica a “Llistar Títols”. El resultat de la consulta apareixerà a la llista interactiva. Si fas click dret a algun títol, tens les mateixes opcions que a la taula de documents.

Un cop tinguis els títols llistats, pots fer doble clic sobre qualsevol títol i se t’obrirà l’editor. En la barra lateral també t’apareixeran les opcions que tens per interactuar amb els documents del resultat de la consulta.

---

<sup>8</sup> Cal assegurar-se que l’autor existeix al sistema



## Llistar autors per prefix

Et permet llistar els autors registrats al sistema tals que el seu nom comenci pel prefix indicat al camp Prefix. S'hi pot accedir a aquest tipus de consulta a través de “Filtre” > “Llistar autors per prefix”.

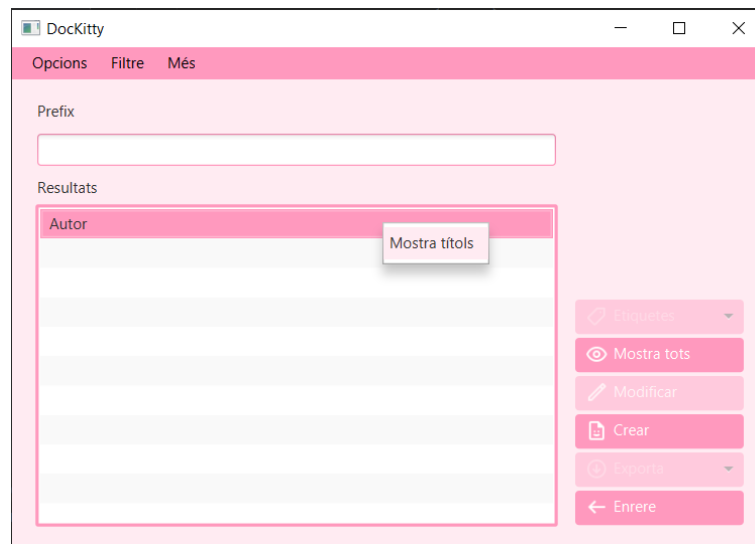


Figura 16. Resultat de llistar autors per prefix

Veuràs la llista de tots els autors del sistema ja que tots els autors comparteixen el prefix buit. A mesura que vagis canviant la informació introduïda al camp de prefix, s'anirà actualitzant la llista d'autors, mostrant més o menys autors.

Pots cercar tots els títols d'aquests autors de dues formes:

- Fent clic dret i clicant a “Mostra títols”
- Fent doble clic a l'autor del que vols saber els títols.

Se't redirigirà a la cerca [Llista títols d'un autor](#) per a veure els documents.

## Cerca k documents semblants a un document

The screenshot shows the DocKitty application window. At the top, there are tabs for 'Opcions', 'Filtre', and 'Més'. Below these, there are input fields for 'Títol' (containing 'títol') and 'Autor' (containing 'autor'). To the right of these fields is a pink button labeled 'K Doc Semblants'. Below the input fields, there is a 'Resultats' section containing a table with three columns: 'Títol', 'Autor', and 'Data'. The first row of the table contains the values 'títol', 'autor', and '2022-12-22 19:56'. To the right of the table, there is a 'K:' input field with the value '1', and two radio buttons labeled 'Bool' (selected) and 'tf-idf'. Below these are several buttons: 'Etiquetes' (with a dropdown arrow), 'Mostra tots', 'Modificar', 'Crear', 'Filtre' (with a dropdown arrow), and 'Enrere' (with a left arrow).

Títol	Autor	Data
títol	autor	2022-12-22 19:56

Figura 17. Resultat de K documents similars

En aquesta opció podràs llistar els k documents més semblants al document seleccionat utilitzant l'estratègia de pesos seleccionada. Aquesta funcionalitat es pot usar de diverses maneres:

- Omplint els camps “Autor”, “Títol” i “K” (enter positiu) manualment.
- Fent clic sobre el document a consultar en comptes d'emplenar els camps manualment, i introduir “K” (enter positiu).
- Des de qualsevol filtre amb taula, fer clic dret sobre el document a la taula o llista, seleccionar l'opció de “k Documents”.

Una vegada escollida qualsevol d'aquestes alternatives, falta escollir una de les dues estratègies de pesos mitjançant uns *radioButtons* amb l'opció “Bool” i “tf-idf”. Finalment, clica el botó de “K Doc Semblants” i el resultat es mostrarà a la taula interactiva.

Cal tenir en compte que el nombre  $k$  ha de ser un natural positiu. En cas contrari, se t'avisarà perquè puguis desfer l'errada.

Per a més informació sobre les estratègies de pesos clica [aquí](#).

Cerca k documents rellevants donades unes paraules

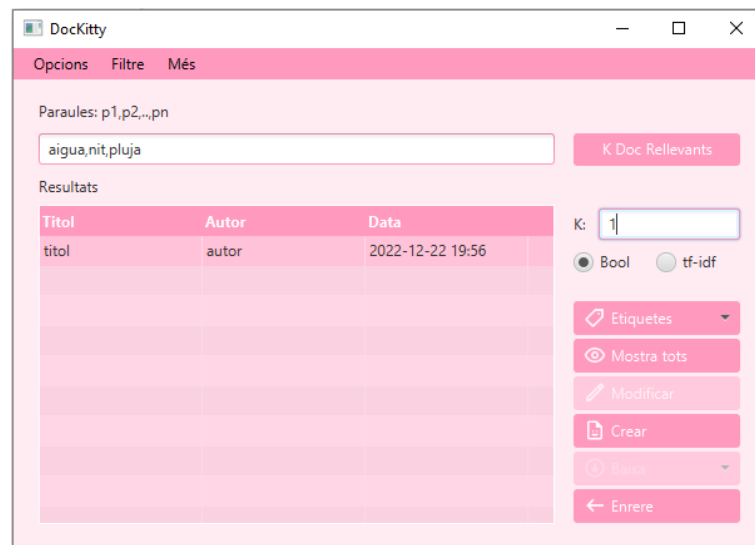


Figura 18. Resultat de K documents rellevants

Funciona de manera molt similar a l'anterior, però ara permet cercar en funció d'un conjunt de paraules que s'introdueixen separades per comes com s'indica a l'aplicació (si no es fa així i es fan servir espais, punts o similars es separaran bé igualment, però recomanem aquesta manera d'entrar la consulta).

Per fer aquesta consulta:

- Introdueix les paraules que vols buscar separades per comes i defineix la "K" (enter positiu).
- Selecciona l'estratègia de pesos a utilitzar "Bool" o "tf-idf" mitjançant els *radioButtons*
- Fes clic sobre el botó "K Doc Rellevants" i el resultat apareixerà a la taula.

Cal tenir en compte que el nombre *k* ha de ser un enter positiu. En cas contrari, se t'avisarà perquè puguis desfer l'errada.

Per a més informació sobre les estratègies de pesos clica [aquí](#).

## Consultes Booleanes

Desde la vista de gestor fitxers és des d'on s'accedeix a consultes booleanes. Per fer-ho, ens dirigim a la barra superior a Filtre > Consultes Booleanes tal com es mostra a la imatge.

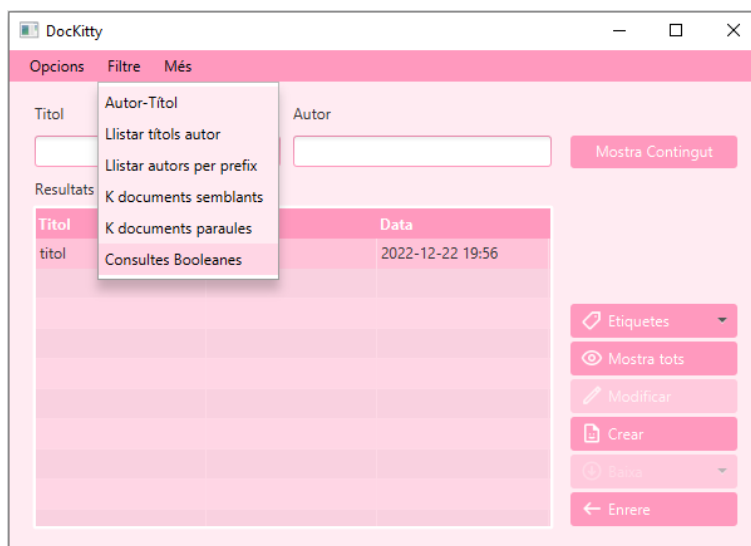


Figura 19. Opció consultes booleanes

## 5. Gestor de Consultes Booleanes

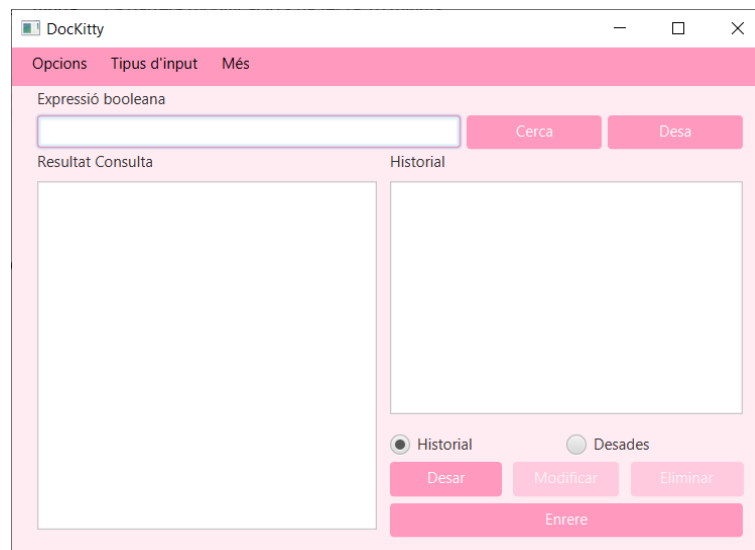


Figura 20. Consultes booleanes per defecte

Aquesta és la pestanya del gestor de consultes booleanes. El primer camp a destacar és el de “Expressió Booleana”, que es el que està per defecte, i permet entrar una expressió per realitzar una cerca.

A sota d'aquest camp, tenim dues llistes: la de l'esquerra, on apareixen els resultats de les consultes (en aquest cas documents) i a la dreta, una altra que serveix tant per mostrar l'historial de consultes booleanes realitzades (és el que apareix per defecte), com per mostrar les consultes booleanes desades. Es mostraran en format:

*Nom (expressió booleana).*

Per canviar entre els dos modes tenim dos *radioButtons* sota la llista per escollir.

L'historial només emmagatzema les 10 últimes consultes fetes, per això tenim les consultes guardades, per desar consultes importants que no volem perdre. En el cas de les consultes desades, ens apareixerà un *popup* per guardar amb el nom que indiquem.

### Tipus d'input

En la barra superior tenim l'opció “Tipus d'input”. Per defecte el tipus d'input és “Expressió Booleana”, que resoldrà una cerca amb l'expressió booleana introduïda.

També podem cercar documents introduint el nom d'una expressió booleana guardada, és a dir, el seu identificador. D'aquesta manera ens estalviem haver d'escriure tota

l'expressió. Per això, però, caldrà desar abans la consulta i assignar-li un nom. Per canviar entre modes d'input, ves al menú superior, i clica "Tipus d'input" i selecciona el que vulguis [Figura 21].

Si fas clic sobre una consulta de l'historial o desada es posarà al camp superior perquè puguis fer la cerca més ràpidament. També pots fer doble clic per realitzar la cerca directament.

Al canviar el tipus d'input entre nom o expressió booleana, quan seleccionem una de les consultes de l'historial o de les desades automàticament s'omplirà el camp de l'expressió amb el nom o la pròpia expressió en funció del cas en el que estiguem.

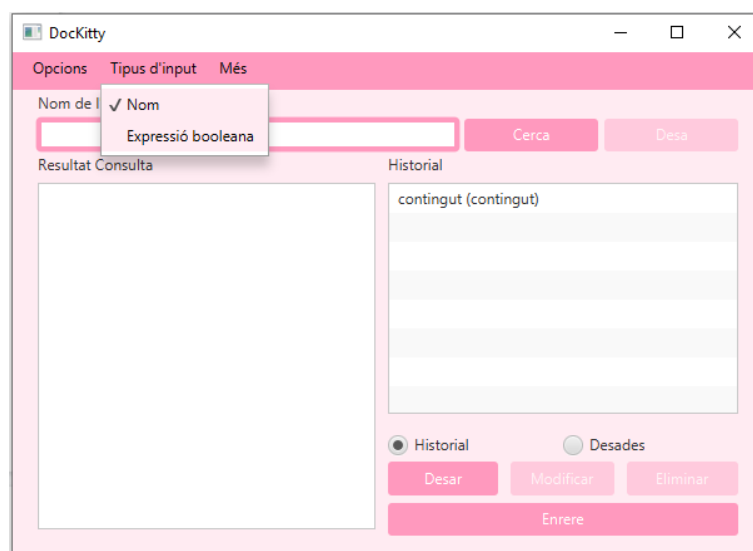


Figura 21. Opcions pel tipus d'input

## Cerca

Per cercar els documents associats a una expressió booleana, cal omplir el camp principal, ja sigui amb una expressió booleana o el nom que la identifica, en el cas d'estar emmagatzemada al sistema, i clicar el botó "Cerca". Es mostrarà el resultat de la consulta com es pot veure a la [Figura 22].

Per a més informació sobre la forma de realitzar cerques clica [aquí](#).

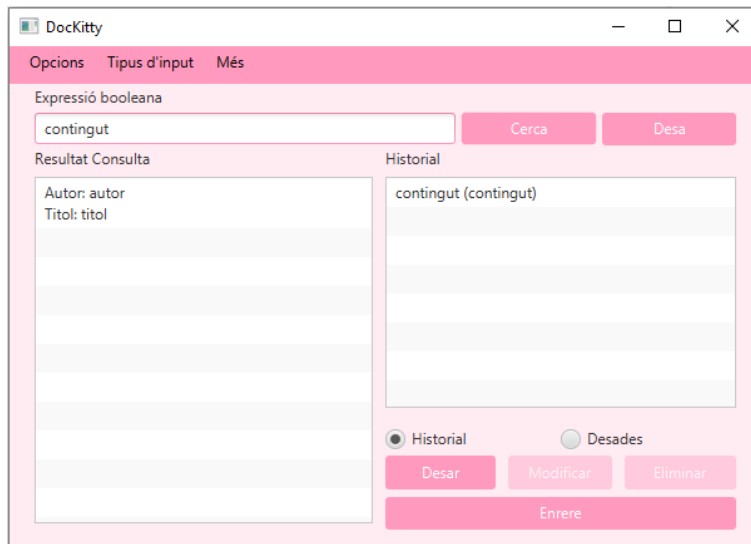


Figura 22. Resultat d'una cerca

Com podem veure, apareixen les consultes que acabem de fer a l'historial, d'aquesta manera si fem doble clic sobre la consulta de l'historial, la podem recuperar per tornar a realitzar la cerca en qualsevol moment com comentàvem abans.

Si fem click dret sobre un dels documents del resultat de la consulta, ens apareixerà un desplegable amb les opcions: “[Mostrar Contingut](#)”, “[K Documents Semblants](#)” i “[Elimina Document](#)”.

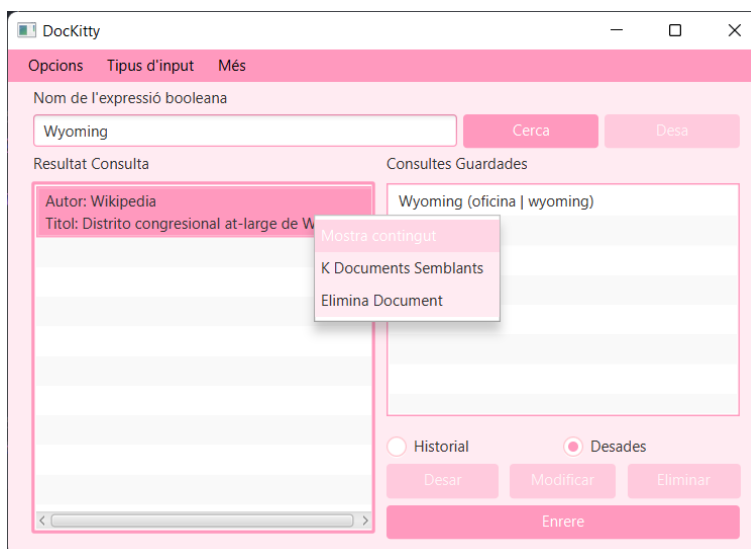


Figura 22.1. Accions disponibles amb el resultat de la consulta

## Desar

Tenim diferents formes de desar una expressió booleana:

- Escrivint-la al camp i clicant al botó de desar de la dreta.
- Realitzant la cerca, seleccionant l'expressió a l'historial i clicant el botó “Desar”.
- Realitzant la cerca, fent clic dret a la llista i clicant a Desar Consulta

Quan desis, apareixerà un *popup* on pots introduir el nom amb el qual vols que s'identifiqui l'expressió.

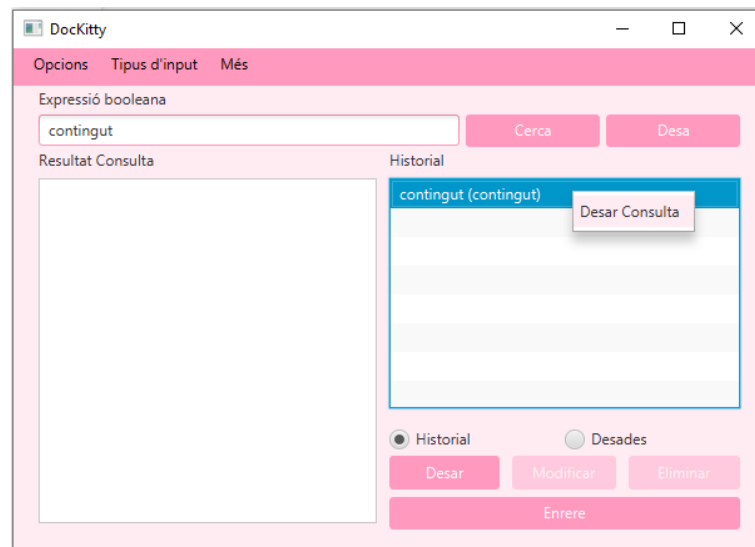


Figura 23. Opció de desar consulta

## Modificar

Si vols modificar el nom d'una consulta selecciona-la a la llista de consultes desades, clica a modificar des del menú desplegable amb el doble clic o amb el botó de “Modificar”, introdueix el nou nom i accepta si vols desar el nou nom o cancel·la si no n'estàs segur.

## Eliminar

Si vols eliminar una consulta desada al sistema, selecciona-la a la llista de consultes desades, clica a “Eliminar” des del menú desplegable amb el doble clic o amb el botó de baix, se't preguntarà si estàs segur de la decisió, accepta si ho estàs o cancel·la en cas contrari.

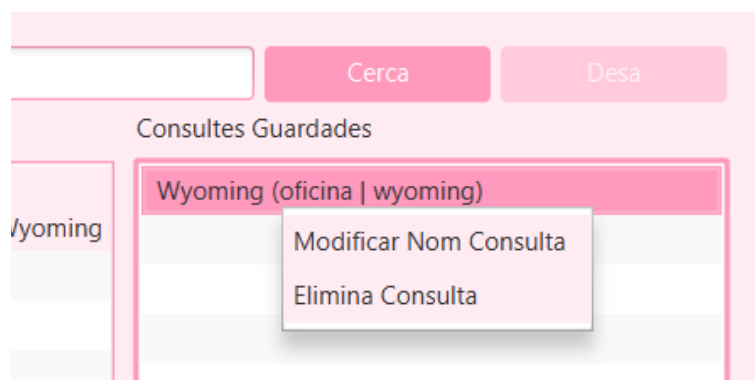


Figura 23.1. Opcions de les consultes desades



## 6. Accions globals

### Exportar Document

Sortirà una pantalla de selecció [Figura 24], on podràs indicar el directori on vols descarregar el fitxer. Com a resultat se't crearà un document amb el nom `titol_autor.format`.

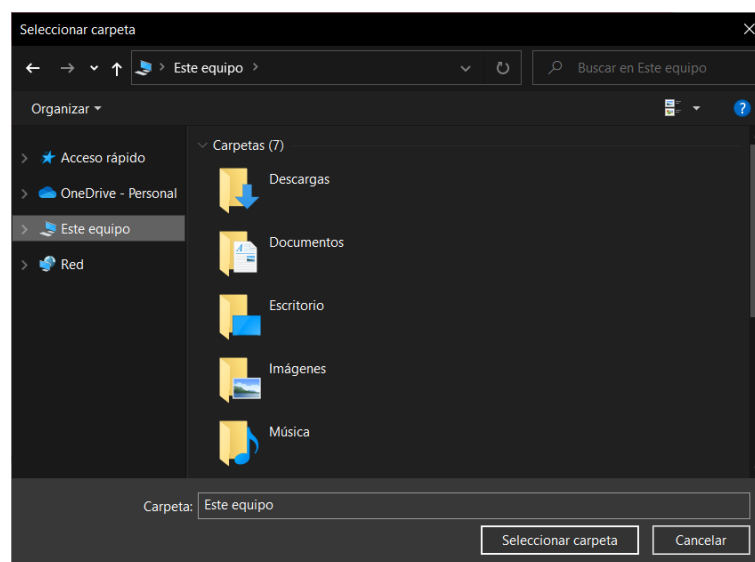


Figura 24. Explorador de fitxers

### Tornar al menú principal i tornar enrere

Per tornar al menú principal, a Fitxer > Menú Principal o Opcions > Menú Principal dependent de la vista. Si es vol tornar a la vista anterior, prémer el botó “Enrere” disponible a cada vista.

### Carregar el nostre repositori

Si estàs interessat en veure el codi d'aquest projecte, pots anar a la barra superior de [Modificar Document](#), [Consultes Booleanes](#) o bé [Gestor Documents](#) i clicar a Més > Gitlab

## 7. Informació adicional

### Funcionalitats generals

Si estàs escrivint text en qualsevol camp on es pugui escriure informació (títol, autor, contingut, expressió booleana...) podràs desfer els últims canvis amb la drecera CTRL+Z, o clicant amb el botó dret del ratolí i seleccionant “Deshacer”.

Si vols refer el que havies desfet, pots utilitzar la drecera CTRL+Y, o clicant amb el botó dret del ratolí i seleccionant “Rehacer”. En la [Figura 25] veiem com és el desplegable i les opcions que ofereix.

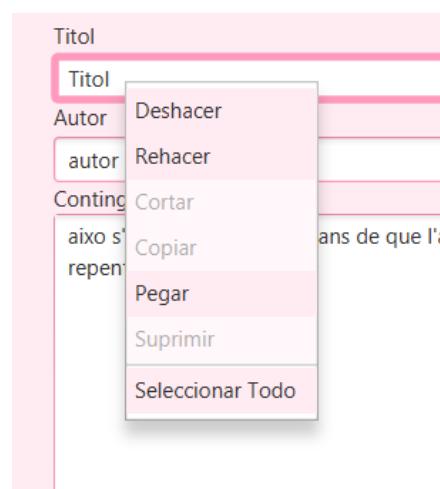


Figura 25. Menú desplegable al fer botó dret a una caixeta on es pot escriure

Per defecte, al tancar el programa donant-li a la X es guarda la sessió per tenir les noves obres, consultes, etiquetes i historials de cara al següent cop que es faci servir l'aplicació. Si per algun motiu es vulgués desar tota la informació abans de tancar l'aplicació (poca bateria a l'ordinador, se'n va la llum, etc) hem agregat l'opció a la barra superior “Opcions” > “Desar sessió”.

Les taules amb contingut que apareixen en les diferents vistes es poden intercanviar d'ordre seleccionant una columna i arrossegant-la cap a la seva nova posició. També es poden redimensionar les columnes.

### Format propietari .kitty

Aquest format es diferencia en que emmagatzema la informació per blocs, s'indica l'inici d'un i fins que no s'indiqui l'inici del següent bloc tota la informació pertanyerà a aquest bloc.

Es diferencia dels altres formats en que l'ordre en que es guardin les diferents parts del document és indiferent, ja que queda totalment encapsulada per cada secció que s'hagi definit. Per indicar les seccions s'utilitza `#nom de la secció#`. Per exemple: `#titol#`.

L'últim bloc del fitxer `.kitty` no estarà delimitat per cap secció posterior (és l'última). El contingut d'aquest darrer bloc serà tota la informació des de l'etiqueta fins al final del document.

Un exemple de document en format `.kitty` seria:

```
#titol#
Això és una mostra
#autor#
grup31.2

#contingut#
Aquí aniria el contingut. Al ser la darrera secció tot el text des d'aquí
fins al final pertanyerà a l'últim bloc. Els salts de línia que hi ha
després de l'autor també s'interpreten com si fos part del bloc, ja que
el bloc s'acaba quan ja no hi ha més text o quan comença la següent
etiqueta.
```

Aquest format permet que la funció de llegir documents pugui llegir les tres etiquetes en qualsevol ordre. No hi ha una limitació del sistema tal i com passa amb el `.txt`. Podem començar escrivint el bloc de contingut o el de títol o el d'autor indiferentment i al parsejar es retornarà l'ArrayList amb la informació a la posició correcta: títol, autor i contingut.

## Estructura de carpetes

Per poder executar el `.jar` del projecte només cal posar al mateix directori on estigui el `.jar` la carpeta "lib" del repositori de Gitlab d'aquesta manera:



Nombre	Fecha de modificación	Tipo	Tamaño
 lib	23/12/2022 21:12	Carpeta de archivos	
 DocKitty.jar	23/12/2022 21:08	Executable Jar File	13.978 KB

Figura 25. Estructura de carpetes

Una vegada fet això obrir una consola al directori i executar aquesta comanda per executar:

Windows:

- `java --module-path .\lib\javafx-sdk-19\lib --add-modules=javafx.controls,javafx.fxml -jar DocKitty.jar`

Per tancar de manera segura l'aplicació, utilitzar sempre la "X" de la barra superior, d'aquesta manera la informació necessària per altres sessions es desarà correctament.

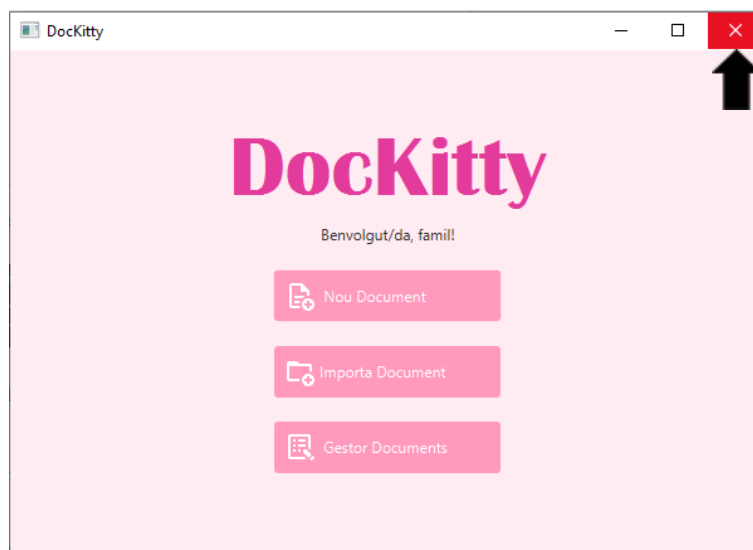


Figura 26. Pàgina principal

## Pesos de l'espai vectorial

Per fer el càlcul de la similaritat entre documents o respecte a un conjunt de paraules, utilitzem la fórmula de la similaritat del cosinus entre els vectors de pesos definida d'aquesta manera:

$$\cos(\theta) = \frac{d \cdot q}{||d|| \cdot ||q||}$$

Aquesta és la fórmula general pel càlcul, per tant, la diferència entre els resultats de similaritat o rellevància depèn de l'assignació de pesos que escollim.

La primera estratègia, es l'assignació booleana de pesos, que assigna un pes de 1 si la paraula està al document o no es té en compte si la paraula no apareix, és a dir té un pes de 0. Aquesta és l'estratègia que s'usa sempre per assignar els pesos al conjunt de paraules consultades en la consulta de “K documents rellevants”, però pel que fa a les pesos de les paraules dels documents, tenim una segona estratègia a escollir, el tf-idf.

Parlem d'aquesta estratègia. Aquesta assignació de pesos té en compte si una paraula és freqüent però també si ho és a través de tots els documents. D'aquesta manera paraules molt freqüents com “cosa” no obtindran tanta rellevància com si ho farien si només tinguéssim en compte la freqüència local a cada document. Aquesta estratègia, busca potenciar la importància de les paraules no tan freqüents que moltes vegades són les que ens permeten diferenciar el contingut de dos textos.

La fórmula per calcular els pesos és aquesta:

$$tf\_idf = tf(t, d) \cdot idf(t, D)$$

- $tf(t, d) = f_{t,d}$ , es pot definir com el nombre de vegades que apareix una paraula al document, i ha diverses maneres de quantificar el tf però aquesta és la més àgil de calcular junt amb la booleana (1 si està 0 si no).
- $idf(t, D) = \log \frac{N}{1 + |\{d \in D: t \in d\}|}$ , es defineix com el logaritme del quocient entre el nombre total de documents i el nombre de documents on apareix t. L'1 del denominador està per si el mot no apareix a cap document.

## Consultes Booleanes

Les consultes booleanes ens permeten cercar aparicions de paraules i/o conjunts de paraules als documents que tenim guardats. Per fer-ho, des de la pantalla de Consultes Booleanes podem introduir una expressió i buscar tots els documents que tinguin almenys una frase que la compleixi.

Podem introduir les consultes de diferents maneres:

- **Introduint una paraula:**
  - Buscarà tots els documents que tinguin aquella paraula en una frase.
  - P. ex: Busquem la paraula *año* i el sistema ens retorna els següents documents, que els dos tenen la paraula año al contingut.

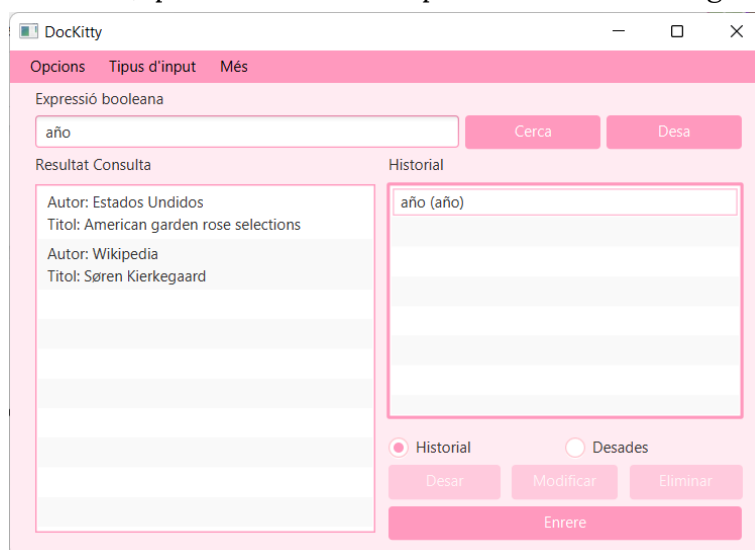


Figura 27. Resultat consulta año

- **Introduint un conjunt de paraules delimitades per { ... }**
  - Buscarà tots els documents que les tinguin totes en una frase.
  - P. ex: Busquem {año 2013 existir} i ens surt el següent document, que conté la frase “Pero, como tantas cosas en los últimos años la economía tuvo su efecto, y desde el **año 2013** en que se otorgaron los últimos galardones el AARS ha dejado de **existir**.”

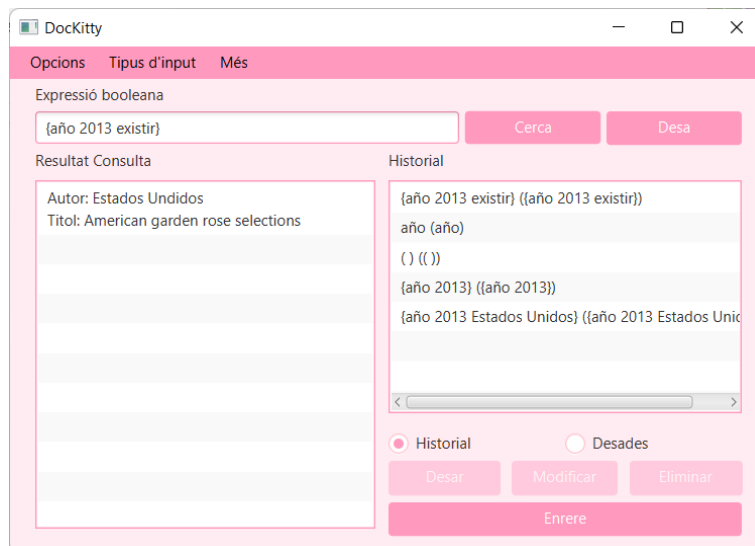


Figura 28. Resultat consulta

- **Introduint una seqüència de caràcters delimitada per “...”**
  - Buscarà els documents que tinguin una frase on aparegui tota aquesta seqüència.
  - P. ex: Busquem “Museo Diocesano de Barcelona” i l’aplicació ens mostra l’únic document que conté la seqüència.

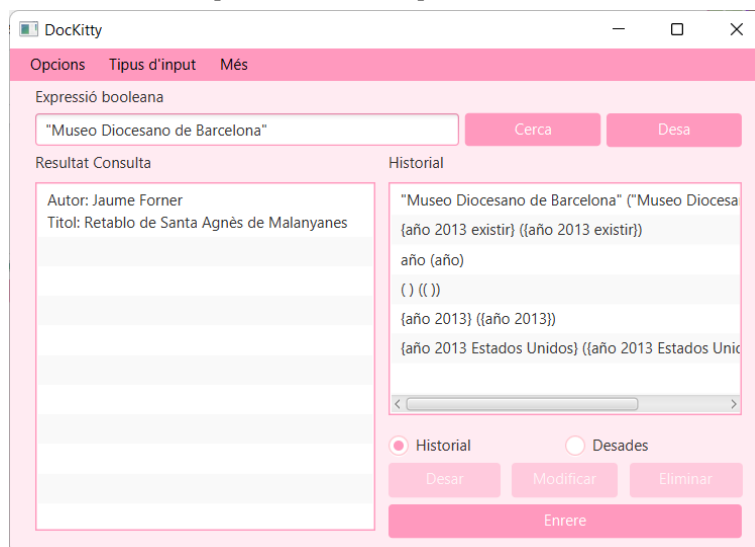


Figura 29. Resultat consulta

- **Encadenant varies consultes**
  - També pots encadenar varies consultes separant-les amb & (i et mostra només els documents que compleixen les dues) o | (et mostra els documents que compleixen una de les dues).
  - Primer es miren les &, i després les | però es pot canviar aquest ordre afegint parèntesis.
  - P. ex: Busquem “pero & (y | o)” i ens retorna els documents que tenen una frase amb “pero” i que tenen una “y” o una “o” a la mateixa frase.



Figura 30. Resultat consulta

- **Negant una expressió**
  - Es pot negar el resultat d'una expressió, és a dir, mostrar tots els documents que no la compleixen posant un ! davant. Si es vol negar una expressió més complexa es pot posar un parèntesis general i negar-lo.
  - P. ex: Si neguem la consulta anterior ens sortiran tota la resta de documents també (I els que ja hi havia també, ja que tenen alguna frase que no ho compleix).

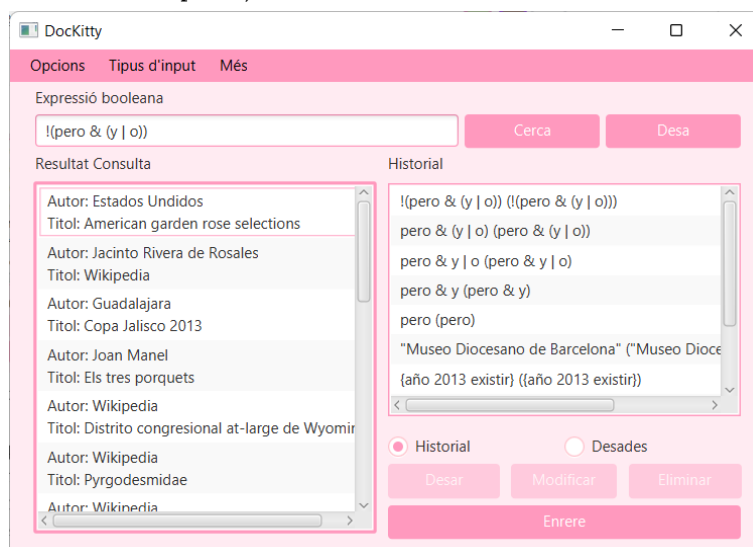


Figura 31. Resultat consulta

Si vols cercar algun dels caràcters especials que s'utilitza per les consultes booleanes es pot buscar posant-lo entre cometes (") i llavors el tractarà com una paraula més