

# ORGANIZÁCIA ELEKTRONICKÝCH DOKUMENTOV

## ŠPECIFIKÁCIA

---

## 1 Popis aplikácie

Cieľom projektu je vytvorenie aplikácie umožňujúcej organizáciu elektronických dokumentov (zatiaľ vo formáte PDF, ak vyjde čas, možno aj v iných formátoch) pre rôznych používateľov.

V prvom rade sa používateľ bude musieť zaregistrovať a prihlásiť. Po prihlásení získa používateľ prístup k svojim dokumentom, ktoré si bude môcť prechádzať, prezeráť si ich priamo v prehliadači, alebo v nich vyhľadávať podľa názvu, obsahu, autora, dátumu, kategórie, tagov, atď.

Samozrejme, používateľ bude musieť najprv dokumenty do svojho účtu pridať. Každý používateľ bude mať na serveri priečinok, do ktorého sa budú dokumenty ukladať, pričom cesta každého dokumentu v danom používateľovom priečniku bude zodpovedať ceste kategórií, do ktorej je dokument zaradený, čo v prípade potreby umožní prehľadné prehľadávanie dokumentov aj na serveri, mimo aplikácie.

Pri pridávaní dokumentu bude môcť používateľ nastaviť metadáta, ako je názov, autor, dátum, kategória, tagy atď. Každý dokument bude môcť byť zaradený iba do jednej kategórie (ktoré budú tvoriť stromovú štruktúru), ale bude k nemu môcť byť priradený ľubovoľný počet tagov. Editovať metadáta bude môcť používateľ samozrejme aj pre už existujúce dokumenty. Bude samozrejme možné aj dokument úplne odstrániť, alebo editovať – tj. reuploadovať nový súbor PDF, ktorý nahradí pôvodný obsah dokumentu, so zachovaním ID, názvu, tagov, oprávnení, atď.

Dôležitou súčasťou aplikácie bude vyhľadávanie podľa obsahu. V prípade, že dokument samotný obsahuje text, aplikácia ho vyextrahuje automaticky, a umožní vyhľadávanie podľa neho. Ak dokument extrahovateľný text neobsahuje (teda ide napríklad o sken), umožní aplikácia používateľovi pridať vlastný textový obsah. V prípade, že bude dosť času, možno sa podarí implementovať aj podporu nejakého OCR.

Používateľ bude môcť ďalej iným používateľom udeľovať oprávnenia k svojim vybraným dokumentom, zatiaľ asi (1) prezeráť, (2) editovať metadáta, (3) reuploadovať samotný dokument, (4) odstrániť.

## 2 Typy používateľov

Aplikácia bude rozoznávať dva typy používateľov:

### **Anonymný používateľ**

Anonymný (neprihlásený) používateľ bude môcť:

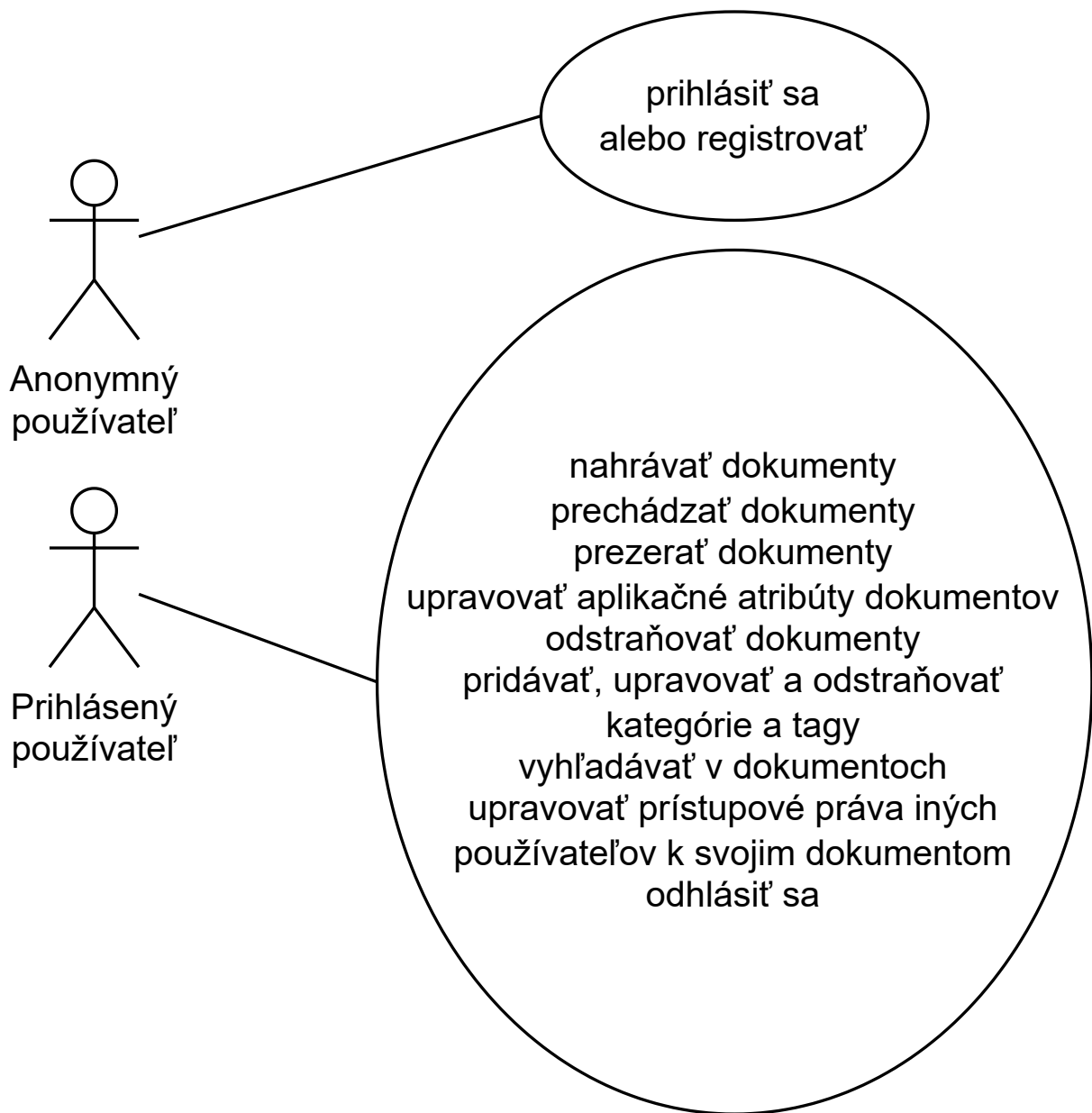
- prihlásiť sa alebo registrovať

### **Prihlásený používateľ**

Prihlásený používateľ bude môcť:

- nahrávať dokumenty
- prechádzať dokumenty
- prezerať dokumenty
- upravovať aplikačné atribúty dokumentov
- odstraňovať dokumenty
- pridávať, upravovať a odstraňovať kategórie a tagy
- vyhľadávať v dokumentoch
- upravovať prístupové práva iných používateľov k svojim dokumentom
- odhlásiť sa

### 3 Use-case diagram



## 4 Popis databázy

Databáza bude obsahovat následující tabulky:

### users

TABLE users		
id	serial	primary key
login	varchar(255)	unique
mail	varchar(255)	unique
password	varchar(255)	not null
name	varchar(255)	
registration_time	timestamp	not null

Stĺpec `password` bude ukladať používateľovo heslo osolené a zahešované pomocou `password_hash/password_verify` PHP API.

### documents

TABLE documents		
id	serial	primary key
owner_id	int	not null
full_path	varchar(255)	not null
name	varchar(255)	not null
author	varchar(255)	
text_content	text	
upload_time	timestamp	not null
creation_time	timestamp	
category_id	int	

### categories

TABLE categories		
id	serial	primary key
parent_id	int	
name	varchar(255)	not null

## tags

TABLE tags		
id	serial	primary key
name	varchar(255)	not null

## rel\_documents\_tags

TABLE rel_documents_tags		
document_id	int	not null
tag_id	int	not null
primary key(document_id, tag_id)		

Reláciu  $m : n$  medzi dokumentami a tagmi implementujeme pomocou špeciálnej tabuľky a dvoch relácií  $1 : n$ .

## permissions

TABLE permissions		
document_id	int	not null
user_id	int	not null
permissions	int	not null
primary key(document_id, user_id)		

Stĺpec `permissions` obsahuje bitovú masku oprávnení, ktoré boli používateľovi s id `user_id` udelené pre dokument s id `document_id`.

## 5 Entito-relačný diagram



## 6 Technologické požiadavky

Na strane servera plánujeme ako databázu používať PostgreSQL a ako skriptovací jazyk PHP (použitie frameworku sa zatiaľ neplánuje). Na strane klienta pravdepodobne ReactJS. Vyvíjať budeme pre prehliadače Mozilla Firefox a Google Chrome, budeme sa snažiť podporovať rozumne aktuálne verzie.

## 7 Časový plán

- 5. týždeň – kostra aplikácie – komponenty, databáza, úvodná stránka, prezeranie zoznamu dokumentov (16 h)
- 6. týždeň – používatelia – registrácia, prihlasovanie, odhlasovanie, ... (16 h)
- 7. týždeň – upload dokumentov, manageovanie adresárovej štruktúry, kategórie (16 h)
- 8. týždeň – prezeranie obsahu dokumentov (10 h), tagy – vytváranie, pridávanie k dokumentom, odoberanie (6 h)
- 9. týždeň – vyhľadávanie – index, vyhľadávací formulár, zobrazovanie výsledkov (16 h)
- 10. týždeň – zdieľanie – udeľovanie a editácia oprávnení iných používateľov k dokumentom (16 h)
- 11. a 12. týždeň – dokončovanie (??? h)