



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

ROČNÍKOVÁ PRÁCE

Dennis Pražák

**Návrh a implementace nástroje na
vytváření diagramů unifikovaných
konceptuálních schémat
multi-modelových a dalších NoSQL
databázových systémů pomocí
prostředků schématických kategorií**

Katedra softwarového inženýrství

Vedoucí bakalářské práce: RNDr. Martin Svoboda, Ph.D.

Studijní program: Informatika

Studijní obor: IPP2

Praha 2021

Poděkování.

Název práce: Návrh a implementace nástroje na vytváření diagramů unifikovaných konceptuálních schémat multi-modelových a dalších NoSQL databázových systémů pomocí prostředků schématických kategorií

Autor: Dennis Pražák

Katedra: Katedra softwarového inženýrství

Vedoucí práce: RNDr. Martin Svoboda, Ph.D., Katedra softwarového inženýrství

Abstrakt: Abstrakt.

Klíčová slova: klíčová slova

Title: Design and implementation of a tool for diagram creation of unified conceptual schemas of multi-model and other NoSQL database systems using schematic categories

Author: Dennis Pražák

Department: Department of Software Engineering

Supervisor: RNDr. Martin Svoboda, Ph.D., Department of Software Engineering

Abstract: Abstract.

Keywords: key words

Obsah

Úvod	2
1 Existující nástroje	3
1.1 Srovnávací kritéria	3
1.2 diagrams.net	4
1.3 drawSQL	7
1.4 ERDPlus	9
Závěr	12
Seznam použité literatury	13
Seznam obrázků	14
Seznam tabulek	15

Úvod

Úvod

1. Existující nástroje

V této kapitole zanalyzujeme některé existující nástroje pro tvorbu diagramů a porovnáme je dle navržených kritérií. Nástroje, které budeme porovnávat jsou

- diagrams.net [1] vhodné pro tvorbu libovolných diagramů,
- drawSQL [2] určené pro tvorbu relačních schémat,
- ERDPlus [3] k vytváření zejména ER diagramů [4].

Před představením existujících nástrojů určíme srovnávací kritéria, dle kterých budeme nástroje analyzovat.

1.1 Srovnávací kritéria

Prvním kritériem pro porovnání nástrojů je jejich kategorie, která vypovídá o účelu nástroje a cílové skupině uživatelů. Základní kategorie jsou

- konceptuální vrstva – tyto nástroje jsou většinou určené pro tvorbu ER diagramů, případně jiným způsobem modelují vztahy a atributy entit, na které při datovém modelování vymezujeme svůj diskurz,
- logická (též technologická) vrstva – tyto nástroje umožňují tvorbu diagramů s ohledem na typ struktur, v kterých jsou data uchovávána, např. relační databáze,
- kresba libovolných diagramů – nástroje, které nejsou omezeny téměř žádným standardem či konvencí a umožňují kresbu libovolných diagramů.

Dalším kritériem je typ úložiště. Nástroje mohou ukládat svá data do paměti prohlížeče (lokálně pro uživatele), na své servery, nebo používat externí úložiště uživatele, například Google Drive¹. Čím více různých typů úložiště nástroj podporuje, tím lépe, neboť uživatel může flexibilně zvolit jeho účelům vyhovující způsob uchovávání dat. Pro interaktivní spolupráci s týmem je lepší sdílené úložiště a pro lokální práci je vhodnější lokální úložiště.

Interaktivní spolupráce je dalším důležitým kritériem. U velkých projektů je vývoj modelu urychlen, pokud nástroj spolupráci umožňuje.

Kromě uložení rozdělané práce do serializovaného formátu musí nástroj umožnit export do formátu, který uživatelé využijí pro své účely. Formáty pro export lze rozdělit do několika kategorií:

- serializovaný formát – většinou se jedná o vlastní formát aplikace a takový soubor nelze jinou aplikací otevřít,
- rastrové formáty, např. PNG² – mají nejširší využití a podporu, lze je použít v dokumentech a na webových stránkách,
- vektorové formáty, např. SVG [5] – nemají tak rozšířenou podporu, nicméně jsou vhodnější v dokumentech po estetické stránce (zvláště při tištění); dále existují vektorové editory, pomocí nichž lze výsledek libovolně upravovat bez

¹<https://www.google.com/drive/>

²Portable Network Graphics – <https://www.w3.org/TR/2003/REC-PNG-20031110/>

potřeby souboru ve serializovaném formátu; většina webových prohlížečů formát SVG podporuje a soubor vykreslí; do této kategorie lze zařadit i jiné otevřené strukturované formáty, např. VSDX³,

- zjednodušený export – některé nástroje šetří práci uživatele tím, že diagram rovnou exportují do HTML⁴, PDF⁵ a podobných finálních formátů pro okamžitou aplikaci, přestože uživatel může zvolit jiný formát a finální vytvořit sám,
- schématické formáty, např. SQL⁶ – téměř výhradně u nástrojů logické vrstvy; umožňují rovnou vytvářet schémata pro databáze.

Stejně jako u typu úložiště, čím více různých formátů exportu nástroj podporuje, tím lépe, neboť nástroj je flexibilní.

Posledním, neméně důležitým kritériem, je způsob komercializace. Většina volně dostupných nástrojů je nějakým způsobem zpoplatněna, ať už se jedná o jednorázový nebo pravidelný poplatek. Nejčastějším komerčním modelem je verze zdarma s omezenými funkcemi a dále několik placených plánů různé úrovně s odemčenými pokročilými funkcemi. U tohoto modelu je důležité vyrovnat funkce tak, aby byl nástroj použitelný i v bezplatné verzi, a aby byly placené funkce atraktivní pro uživatele. Při srovnávání budeme věnovat pozornost i tomu, jestli jsou placené funkce esenciální.

1.2 diagrams.net

Srovnávací kritéria:

- kategorie – kresba libovolných diagramů,
- typ úložiště – lokální, externí, prohlížeč,
- export – serializovaný, rastrový, vektorový, zjednodušený,
- interaktivní spolupráce – částečně podporována (pomocí externích úložišť),
- komercializace – veškeré funkce jsou zdarma a není potřeba uživatelský účet; z jiného pohledu lze počítat cenu externích úložišť, ale ta jsou volitelná.

Nástroj diagrams.net [1], dříve draw.io, je obecný open-source kreslicí nástroj (který však nepřijímá změny od externích vývojářů) vydaný s licencí Apache License 2.0⁷, dostupný jako webová aplikace⁸ nebo jako desktopová aplikace. Desktopová verze aplikace je sestavena stejným způsobem jako webová, pouze je zabalena pomocí platformy Electron [6] do okna Chromium. Je vyvinut v běžných webových technologiích (JavaScript⁹, CSS¹⁰, HTML).

³Microsoft Visio XML formát založený na ISO 29500 – <https://interoperability.blob.core.windows.net/files/MS-VSDX/%5bMS-VSDX%5d.pdf>

⁴HyperText Markup Language – <https://www.w3.org/TR/2021/SPSD-html52-20210128/>

⁵Portable Document Format, ISO 32000 – <https://www.iso.org/standard/75839.html>

⁶Structured Query Language – <https://www.iso.org/standard/63555.html>

⁷<https://www.apache.org/licenses/LICENSE-2.0>

⁸na adrese <https://app.diagrams.net>

⁹Standardizován jako ECMAScript, ISO 16262 – <https://www.iso.org/standard/55755.html>

¹⁰Cascading Style Sheets – <https://www.w3.org/TR/css/>

Diagramy lze uložit do serializovaného XML¹¹ formátu .drawio. V tomto formátu je pro každý diagram XML element **diagram**, ve kterém se nachází data zakódována do Base64¹². Tato data jsou komprimována pomocí zlib¹³ a obsahují další XML dokument (URL-encoded¹⁴, tj. zakódovaný), tentokrát již serializaci vlastního diagramu. Formát tak není bez dekomprese čitelný člověkem. Výhodou je, že lze uložit více diagramů do jednoho souboru a každý pojmenovat. Rozhraní k tomu určené je identické s listy souboru tabulkových procesorů, jako Microsoft Excel¹⁵ a Google Sheets¹⁶.

Soubor s diagramy lze také uložit do formátu SVG, který je navíc otevřený a podporují ho jiné nástroje. Uživatel má při exportu k dispozici možnost *Include a copy of my diagram*, která do SVG souboru zahrne již zmíněný Base64 řetězec, ve kterém je diagram serializovaný. Ve výsledku to znamená, že takto exportované SVG soubory umí diagrams.net i otevřít a práce na nich může plnohodnotně pokračovat. Toto řešení se nám líbí, protože se jedná o schování vlastního formátu do SVG, který je nejvhodnějším pro přechovávání a zobrazování diagramů.

Dalšími možnostmi exportu a ukládání jsou

- rastrové soubory PNG, JPEG¹⁷,
- soubor PDF, do kterého je ve vektorovém formátu diagram vložen,
- soubor HTML, do kterého lze podobně jako v SVG data diagramu uložit v serializované formě, případně pouze vložit veřejný odkaz URL na diagram (pokud je použito odpovídající úložiště); v tomto souboru je pak zahrnut JavaScript od diagrams.net, který diagram vykreslí,
- otevřený formát VSDX, původně vyvinutý pro Microsoft Visio.

Ze stejných souborů lze diagramy také importovat, ovšem editovat je lze jen pokud je v nich zahrnut formát drawio, čehož je dosaženo u některých formátů popsaných výše.

Jako úložiště si lze vybrat Google Drive, OneDrive¹⁸, Dropbox¹⁹, GitHub²⁰, GitLab²¹, paměť prohlížeče a místní úložiště (disk uživatele). Soubor lze ze stejných úložišť i otevřít a importovat, navíc k tomu i z libovolné dostupné URL.

Interaktivní spolupráce je umožněna pouze pokud soubor jako úložiště využívá takové, ke kterému mají přístup zápisu (popř. pouze čtení) všichni účastníci se uživatelé (Google Drive, OneDrive, Dropbox, GitHub, GitLab). Tato úložiště je však nutno manuálně vhodně nastavit (přístup ostatním uživatelům). U všech úložišť je rychlost reflektování změn ostatních uživatelů podobná – vcelku pomalá, protože aplikace musí změny aktivně kontrolovat a načítat.

¹¹Extensible Markup Language – <https://www.w3.org/TR/xml/>

¹²RFC 2045 §6.8 – <https://datatracker.ietf.org/doc/html/rfc2045#section-6.8>

¹³<https://zlib.net/>

¹⁴RFC 3986 §2.1 – <https://datatracker.ietf.org/doc/html/rfc3986#section-2.1>

¹⁵<https://aka.ms/excel>

¹⁶<https://sheets.google.com>

¹⁷Joint Photographic Experts Group, ISO 19566 – <https://www.iso.org/standard/65348.html>

¹⁸<https://aka.ms/onedrive>

¹⁹<https://dropbox.com>

²⁰<https://github.com/>

²¹<https://gitlab.com>

Menu **File** → **Publish** chybně napovídá, že se jedná o funkci interaktivní spolupráce. Ve skutečnosti je uživateli jen zobrazen odkaz na soubor ve vybraném úložišti (ale pouze pro Google Drive a OneDrive, jinak je tato možnost vypnuta). Spolupracující uživatel tak musí tento soubor v daném úložišti uložit k sobě (sdíleně), aby mohla spolupráce začít.

Jako další možnost jsme zvažovali desktopovou aplikaci s načteným souborem, který je libovolným externím nástrojem sdílen mezi uživateli. Bohužel, soubor se nepřenačítá automaticky, ale musí být manuálně synchronizován tlačítkem **File** → **Synchronize** (**Alt+Shift+S**), které je dostupné pouze v desktopové verzi aplikace. Uživatel je při externí změně souboru upozorněn (avšak ne spolehlivě vždy) červeným nápisem. Algoritmus synchronizace funguje správně a tak, jak uživatel očekává.

Nejlepší způsob dosažení interaktivní spolupráce je dle našeho názoru volba systému pro správu Git²² repozitářů (GitLab nebo GitHub), protože

1. tato úložiště jsou dostupná jak z webové, tak z desktopové verze aplikace,
2. synchronizace probíhá pomocí systému Git,
3. díky použití systému Git lze jednoduše spravovat verze a body v historii při vývoji diagramu.

K poslednímu bodu je třeba podotknout, že jiná webová úložiště také podporují správu verzí, avšak není tak rozvinutá, jako správa systémem k tomu určeným – Git. Diagrams.net sám o sobě správu verzí neobsahuje, jen obvyklé “Undo, Redo” pro aktuálního uživatele. Úpravy ostatních uživatelů nelze vracet postupně, lze se pouze vrátit za bod synchronizace.

Uživateli jsou v levém postranním panelu k dispozici standardní tvary ER diagramů, UML diagramů [7], flowchart diagramů a další základní tvary pro kresbu diagramů. Tvary lze libovolně kombinovat a spojovat podržením levého tlačítka a tažením myši z a do kotev na krajích objektů. Každý objekt a spojovací čára má vlastnosti, které lze upravovat v pravém postranním panelu. Upravovat lze přímo i vlastnosti formátu SVG.

Uživatelské rozhraní, které je vidět na obrázku 1.1, je velmi podobné kancelářskému aplikacím Google. Je tak přívětivé pro nové uživatele, kteří již s aplikacemi Google dříve pracovali.

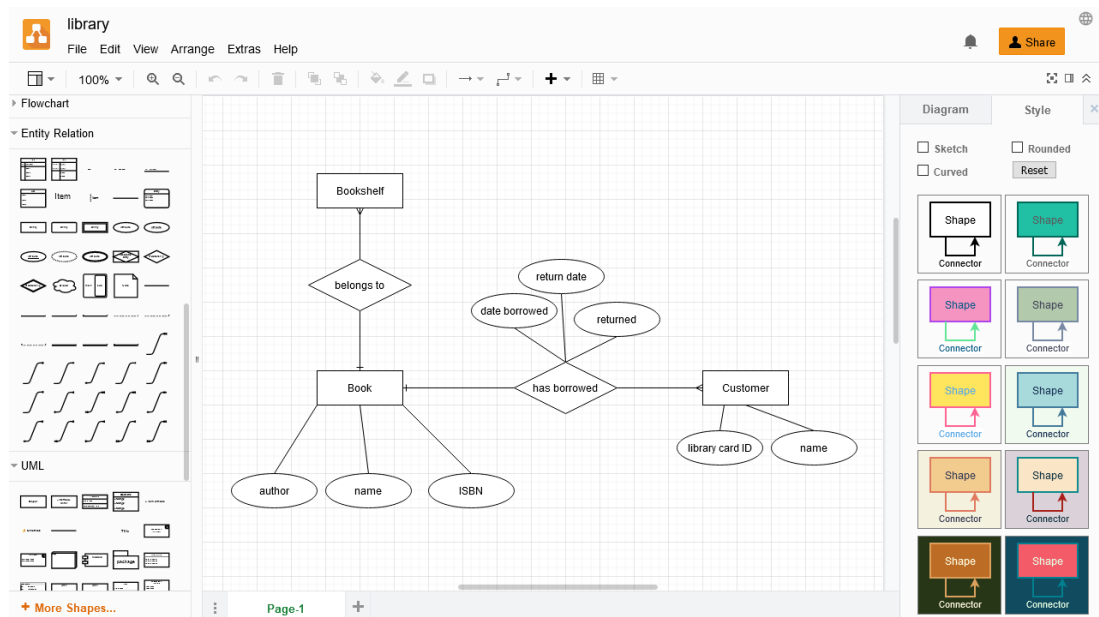
Jako výhody určujeme

- univerzálnost a flexibilita – nástroj lze použít pro tvorbu jakýchkoli diagramů,
- množství podporovaných formátů – export pokrývá téměř všechny možné účely,
- cena – všechny funkce jsou zdarma,
- více diagramů v jednom souboru

a nevýhodami jsou

- chybějící možnost pro export do (jednoduše) strojově zpracovatelného formátu, nelze tak bez lidské práce diagram převést do logické vrstvy (to je zapříčiněno obecností nástroje, jeho účelem je kresba, ne abstrakce),

²²Systém pro správu verzí Git – <https://git-scm.com/>



Obrázek 1.1: Tvorba ER diagramu v aplikaci diagrams.net

- pomalé zobrazování změn při interaktivní spolupráci, zároveň není zpočátku jasné, jak spolupráce dosáhnout.

Výhodou i nevýhodou může být nutnost použití externího úložiště. Pro velké společnosti se může jednat o bezpečnostní opatření, protože diagrams.net k diagramům nemá přístup. Pro malé týmy se může jednat o nevýhodu, protože je potřeba účet na externím webu, nebo jiný způsob sdílení a správa tohoto úložiště.

1.3 drawSQL

Srovnávací kritéria:

- kategorie – logická vrstva,
- typ úložiště – online, poskytované autory produktu
- export – schématický (obecný SQL i platformě specifické formáty), rastrový PNG, serializovaný (JSON [8], v době psaní práce se chystá)
- interaktivní spolupráce – pouze v placené verzi,
- komercializace – omezená verze navždy zdarma, různé měsíčně placené plány.

Nástroj drawSQL [2] je modelovací nástroj pro tvorbu relačních schémat. Aplikace je dostupná ve webovém prohlížeči²³. Je vyvinuta ve standardních webových technologiích a používá framework Vue.js. Plán zdarma umožňuje tvorbu veřejně přístupných diagramů, které mohou mít maximálně 15 tabulek (entit). Měsíčně placené plány umožňují vytvářet neveřejné diagramy, více (až neomezeně mnoho) tabulek v diagramu, více uživatelů, kteří mohou na diagramu spolupracovat, a přístup k verzovacím nástrojům. K vyzkoušení i používání nástroje je potřeba uživatelský účet.

²³na adrese <https://drawsql.app>

Hlavní funkcí drawSQL je export schématu do SQL. Proto si uživatel při vytváření diagramu zvolí cílovou databázi, pro kterou schéma tvoří. Výsledné SQL tak bude mít tvar, se kterou cílová databáze umí pracovat. Podporovanými databázemi jsou MySQL²⁴, PostgreSQL²⁵ a SQL Server²⁶.

Rozhraní, které je vidět na obrázku 1.2, obsahuje diagram a postranní panel. V postranním panelu lze vytvářet jednotlivé tabulky, definovat jejich sloupce a vlastnosti jednotlivých sloupců – typ sloupce, nullability²⁷, zda se jedná o primární klíč, unikátní klíč nebo index. Tyto změny se v reálném čase reflektují v diagramu, ve kterém může uživatel jednotlivé sloupce spojovat, čímž vytváří cizí klíče. Pozici těchto lomených čar lze upravovat pouze posunutím tabulky v diagramu. Pokud je cizích klíčů víc, začne být diagram velmi nepřehledný.

Diagram lze importovat ze souboru SQL stisknutím **File** → **Import**. Stisknutím tlačítka **File** → **Export** se otevře nabídka Export, ve které může uživatel diagram exportovat do SQL své předem zvolené databáze, nebo do rastrového obrázku ve formátu PNG. Vývojáři aplikace plánují implementovat také export diagramu pomocí serializace do formátu JSON. V nabídce Export je navíc možnost nechat si vygenerovat platformně specifický kód jako například migrační třídy pro Laravel²⁸, definice modelů pro Laravel a migrační schémata pro AdonisJS²⁹.

Interaktivní spolupráce je k dispozici pouze v placené verzi. Dle našeho názoru je interaktivní spolupráce hlavní funkcí tohoto nástroje oproti konkurenčním relačním modelovacím nástrojům. Některá integrovaná vývojová prostředí (např. Visual Studio³⁰) obsahují nástroj pro relační modelování i generování databázového schématu. Hlavním omezením těchto nástrojů je však absence interaktivní spolupráce, jedná se spíše o spolupráci iterací. Proto považujeme určení interaktivní spolupráce za placenou funkci za negativní rozhodnutí pro využitelnost nástroje v relaci s konkurencí.

Web drawSQL také zveřejňuje šablony modelů³¹ (jedná se spíše o příklady). Šablony jsou většinou potenciální modely známých produktů (např. WordPress³²) a tvoří je autoři drawSQL. Tuto funkci považujeme za výhodu, protože společnosti a individuální vývojáři se mohou inspirovat existujícími a ověřenými řešeními, případně nezačínat se svým modelem od nuly.

Závěrem určíme výhody drawSQL:

- příjemné uživatelské rozhraní (viz obrázek 1.2),
- možnost určení typu relace, o sémantiku se aplikace stará sama (one-to-one, one-to-many, many-to-many),
- několik platformně specifických generátorů modelu,
- šablony a příklady existujících modelů

²⁴<https://mysql.com/>

²⁵<https://postgresql.org/>

²⁶Microsoft SQL Server – <https://aka.ms/sqlserver>

²⁷*nullability* je příznak, který určuje, zda lze sloupec v řádku nastavit na hodnotu NULL

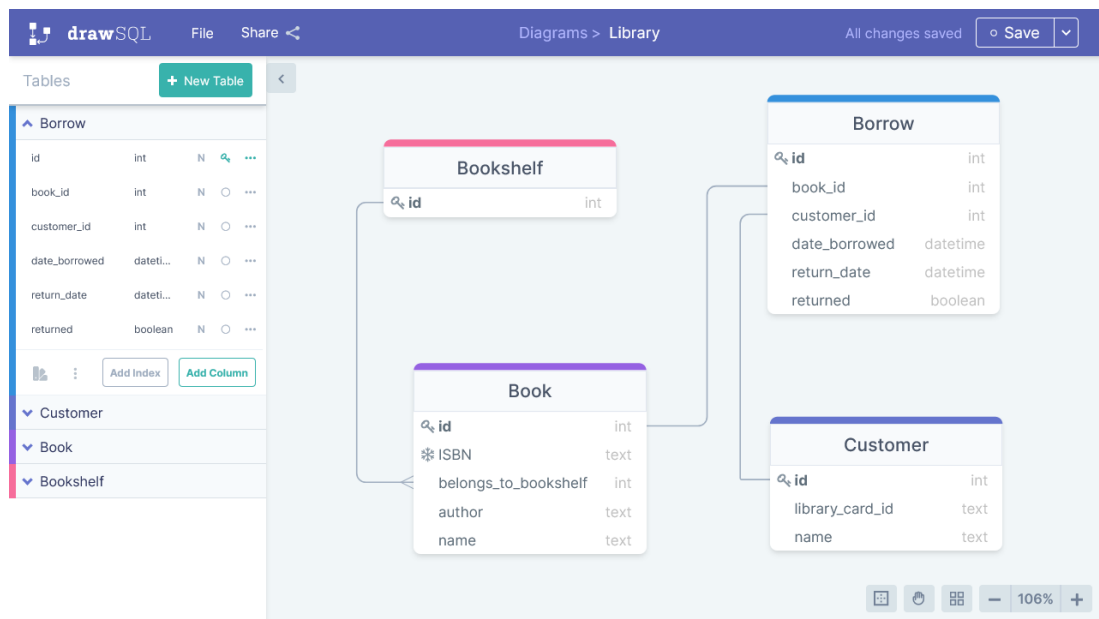
²⁸Framework pro PHP – <https://laravel.com/>

²⁹Framework pro Node.js – <https://adonisjs.com/>

³⁰Vývojové prostředí Microsoft Visual Studio – <https://visualstudio.microsoft.com/>

³¹na adrese <https://drawsql.app/templates>

³²<https://wordpress.com/>



Obrázek 1.2: Tvorba diagramu v drawSQL

a nevýhody:

- nelze upravit ani přesunout lomené čáry spojující cizí klíče, což způsobuje chaos pokud je v diagramu větší množství entit,
- interaktivní spolupráce pouze v placeném plánu,
- správa verzí pouze v placeném plánu,
- k vyzkoušení nástroje je potřeba uživatelský účet,
- podporuje pouze relační databáze.

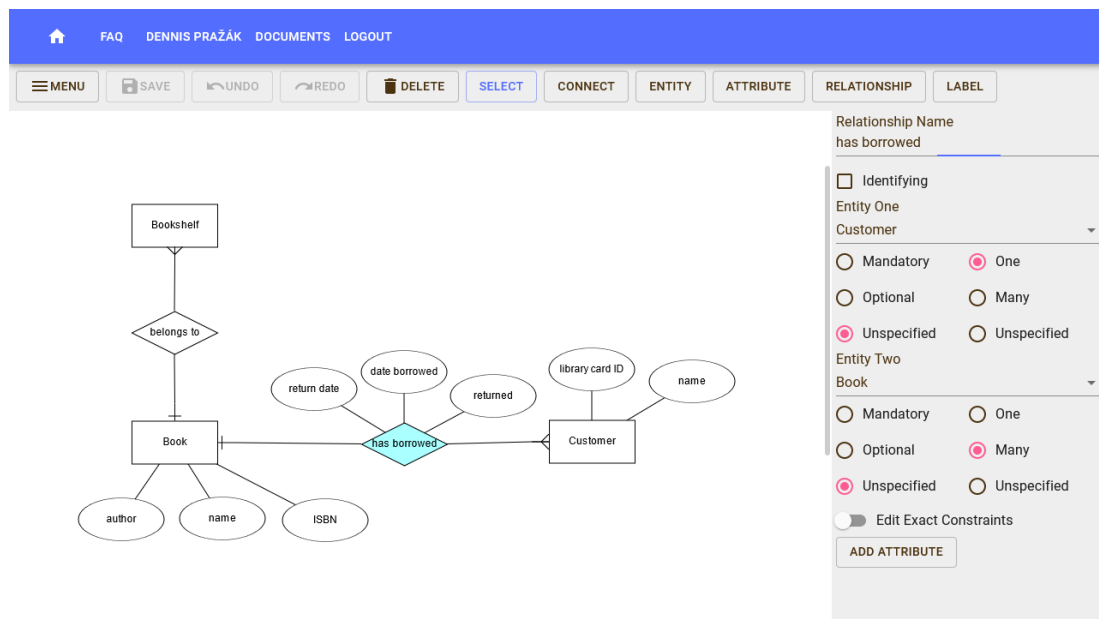
1.4 ERDPlus

Nástroj ERDPlus [3] je modelovací nástroj pro tvorbu ER diagramů, relačních schémat a hvězdicových schémat. Aplikace je dostupná ve webovém prohlížeči³³. Její uživatelské rozhraní je tedy vyvinuto ve standardních webových technologiích – HTML, CSS a JavaScript – a dále využívá framework React [9] pro tvorbu rozhraní v jazyce JavaScript.

ERDPlus lze používat bez založení uživatelského účtu a vytvořený diagram exportovat do speciálního formátu erdplus, nicméně uživatel tak přijde o možnost využití úložiště diagramů na serveru aplikace. Služby ERDPlus nejsou žádným způsobem zpoplatněny.

Tvorba ER diagramů je intuitivní s jednoduchým uživatelským rozhraním, které je vidět na obrázku 1.3. Uživatel má na výběr mezi vytvořením entity, atributu, relace, spojení mezi těmito objekty a jednoduchého textového popisku. V pravé části rozhraní se nachází panel s vlastnostmi zvoleného objektu. V tomto panelu může uživatel také rychleji tvořit atributy entit a relací. Při zvolení relace

³³na adrese <https://erdplus.com/>



Obrázek 1.3: Tvorba ER diagramu v ERDplus

lze v panelu zvolit entity, které mají být v relaci, a spojení je pak automaticky vytvořeno. Zároveň lze zvolit jednotlivé multiplicity relace.

ER diagram lze exportovat do rastrového formátu PNG. Zajímavou funkcí je také převod do relačního schématu. Tato funkce je dostupná pouze tehdy, když uživatel ER diagram uloží na server ERDPlus. Poté zvolí možnost *Convert to Relational Schema* a ERDPlus vytvoří nové relační schéma. Z relačních schémat lze podobně vygenerovat SQL.

název produktu	diagrams.net	drawSQL	ERDPlus
kategorie (vrstva)	libovolné diagramy	logická	konceptuální
serializovaný export	✓	✗ ^a	?
rastrový export	✓	✓	✓
vektorový export	✓	✗	✗
schématický export (SQL)	✗	✓	✓
zjednodušený export	HTML, PDF	platformně- specifický scaffolding	?
poskytuje úložiště	✗ ^b	✓	✓
využívá paměť prohlížeče	✓		

a. plánovaná funkce

b. využívá úložiště třetích stran

Tabulka 1.1: Srovnání existujících řešení

Závěr

Závěr

Seznam použité literatury

- [1] JGraph Ltd. diagrams.net. [online]. URL: <https://www.diagrams.net/>.
- [2] DrawSQL. DrawSQL – Database Schema Diagrams. [online]. URL: <https://drawsql.app/>.
- [3] ERDPlus. ERDPlus. [online], 2021. URL: <https://erdplus.com/>.
- [4] Peter Pin-Shan Chen. The entity-relationship model—toward a unified view of data. *ACM Trans. Database Syst.*, 1(1):9–36, březen 1976. doi:10.1145/320434.320440.
- [5] Schulze Dirk and SVG Working Group. Scalable Vector Graphics (SVG) 2 Specification. W3C Recommendation CR-SVG2-20181004, W3C, říjen 2018. URL: <https://www.w3.org/TR/2018/CR-SVG2-20181004/>.
- [6] OpenJS Foundation. Electron. [online], 2021. URL: <https://www.electronjs.org/>.
- [7] Object Management Group (OMG). Unified Modeling Language version 2.5.1. [online], prosinec 2017. URL: <https://www.omg.org/spec/UML/>.
- [8] TC39 Group. The JSON data interchange syntax 2nd Edition. Standard ECMA-404, ECMA International, Rue du Rhône 114, CH-1204 Ženeva, Švýcarsko, prosinec 2017. URL: <https://www.ecma-international.org/publications-and-standards/standards/ecma-404/>.
- [9] Facebook Inc. React, A JavaScript library for building user interfaces. [online], 2021. URL: <https://reactjs.org/>.

Seznam obrázků

1.1	Tvorba ER diagramu v aplikaci diagrams.net	7
1.2	Tvorba diagramu v drawSQL	9
1.3	Tvorba ER diagramu v ERDplus	10

Seznam tabulek

1.1 Srovnání existujících řešení	11
--	----