

LIBRERA SEARCH SOLUTION

Enables indexing content from PDFs making it searchable
with the possibility to classify them by the Web UI

Contents

Requirements.....	2
Overview.....	3
Clean Architecture.....	4
Scope.....	5
RQ-01 Automated Task Service.....	6
RQ-02 Indexer.....	9
RQ-03 Model Layer.....	10
RQ-04 Web API.....	11
RQ-05 Front end App.....	13
RQ-06 Database.....	17
Deployment.....	18
Software life cycle management.....	19
Annexes.....	20

Doc. Version	Date	Author	Description
1.0	2025-12-17	https://github.com/kraftcoding	First version of the release

Requirements

The next table shows the requirements for the solution described.

Requirement	Description
RQ-001 Automated Task Service	Hangfire offers the possibility to manage automated task services to process PDFs
RQ-002 Indexer	The indexer libraries that contains the logic to process the PDFs
RQ-003 Model Layer	The model layer with EF core to interact with database
RQ-004 Web API	The back-end to execute CRUD operations in server
RQ-005 Front end App	Angular App to offer UI to manage the records and associated records
RQ-006 Database	The data model, full text search and index services and the logic needed by Stored Procedures

Overview

This document contains the technical details of the Librera Search v1.x solution built with Angular 16, Hangfire Server and .Net 8; enables to index content from PDFs making it searchable with the possibility to classify them by the Web UI. Implements token-based authentication for registered users and policy-based authorization in the endpoints.

The application contains the next functionality:

- Each register has id, title ,authors ,series, ids, published, publisher, languages, tags, formats, path and indexed content
- We can create, retrieve, update, delete records, launch the index service and search PDF content linked to the records
- Includes two pages for finding registers by title or search indexed content
- JWT (JSON Web Token) format used for authentication
- Policy-based authorization in the endpoints

Token-based authentication is a popular approach to securing web applications, providing a stateless and scalable way to verify user identities. In this implementation, we'll explore how to integrate token-based authentication using Angular 16 as the frontend framework and .Net 8 WebAPI as the backend API.

Token-based authentication involves generating a unique token upon successful user authentication, which is then passed with each subsequent request to verify the user's identity. This approach eliminates the need for server-side session storage, making it ideal for modern web applications.

Key Components:

- .Net 8 WebAPI: Handles token generation and validation
- Angular 16: Manages user authentication and includes the token in requests
- JWT (JSON Web Token): The token format used for authentication
- Clean architecture

This implementation provides a secure and efficient way to authenticate users, allowing for seamless interaction between the frontend and backend.

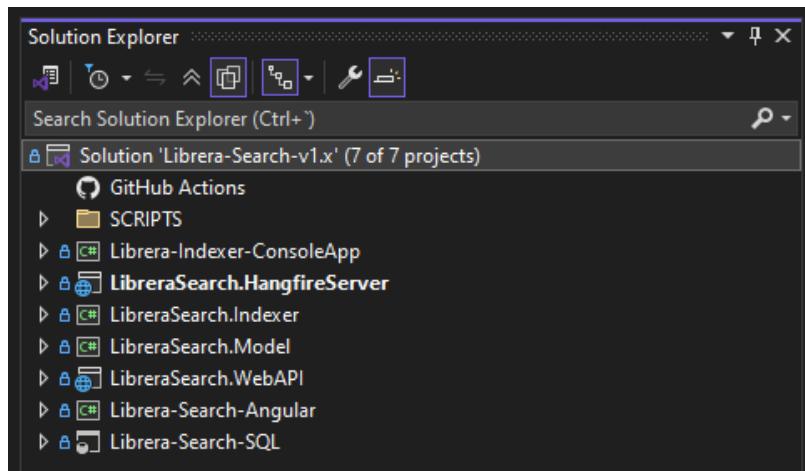
In ASP.NET, policies are used to define requirements that must be met in order to give a user the authorization to access an endpoint.

Within Web API projects you can use policies to guard and protect endpoints from unauthorized users that don't meet the criteria. Instead of writing the authorization logic in the endpoint itself, you can refactor this into a policy that is reusable across multiple endpoints. By extracting the authorization logic from the endpoint, you can make your code more clear and keep your endpoints concise.

Clean Architecture

Clean Architecture is an architecture pattern that is based on four layers: presentation, application, domain, and infrastructure. The presentation layer is responsible for handling user interactions, the application layer is responsible for handling the business logic, the domain layer is responsible for defining the business entities and their relationships, and the infrastructure layer is responsible for handling external concerns such as databases and web services.

Here the solution structure (Visual Studio 2022):



The key principles of Clean Architecture include the independence of the layers, the data flow from the outer layers to the inner layers, and the ability to swap out dependencies without affecting the rest of the system. Clean Architecture emphasizes the separation of concerns to make the codebase easier to maintain, test, and extend.

The project structure of Clean Architecture follows a structured layout that enables code reusability and testability. In the project structure, you would typically have a project for the presentation layer, a project for the application layer, a project for the domain layer, a project for the infrastructure layer, and a project for the tests. Each layer would have its own set of interfaces and implementations, with the application layer acting as the bridge between the other layers.

Scope

This first version of the solutions grants a view of how could be easily implemented in .NET 8 a simple but effective tool to manage book records and make content searchable using open source elements like Hangfire and front-end technologies like Angular.

RQ-01 | Automated Task Service

An easy way to perform background processing in .NET and .NET Core applications. No Windows Service or separate process required. Backed by persistent storage. Open and free for commercial use.

Here the Book Indexer Job definition in the dashboard:

The screenshot shows the Hangfire Dashboard for the 'BookIndexerJob'. The top navigation bar includes links for 'Hangfire Dashboard', 'Jobs (1)', 'Retries (0)', 'Recurring Jobs (1)', 'Servers (1)', and a 'Back to site' link. The main content area is titled 'BookIndexerJob' and displays the following information:

- Job Status:** The job is finished. It will be removed automatically *in a day*.
- Code Snippet:**

```
// Id: #15
using HangfireExample.Jobs;

var bookIndexerJob = Activate<BookIndexerJob>();
await bookIndexerJob.RunAsync();
```
- Parameters:**

CurrentCulture	"en-US"
CurrentUICulture	"en-US"
RecurringJobId	"BookIndexerJob"
Time	1764923111
- State:**

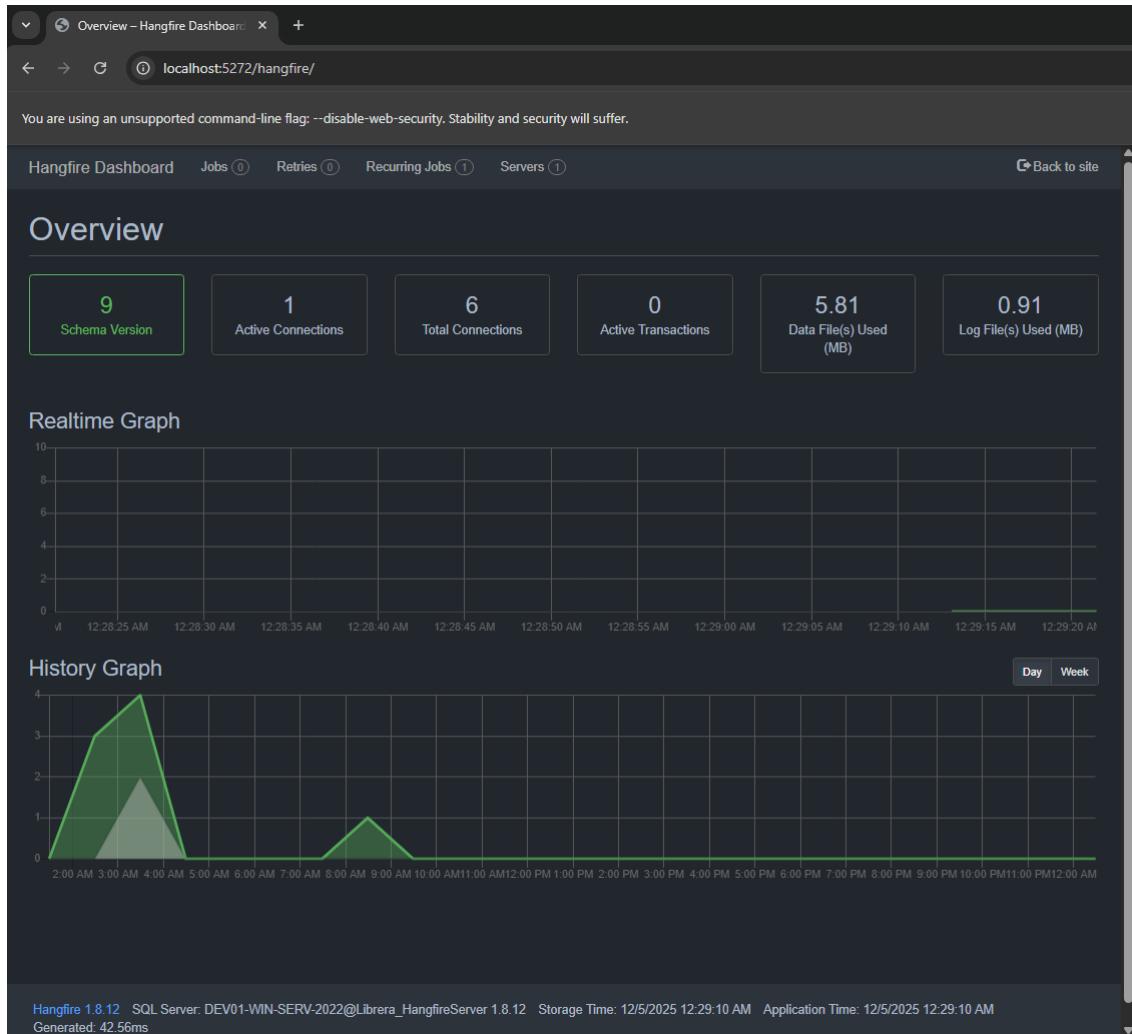
Succeeded	6 minutes ago (+39.010s)
Latency:	9.034s
Duration:	38.989s

Processing	+6.970s
Server:	DEV01-WIN-SERV-10292
Worker:	0955c009

Enqueued	+2.053s
Triggered by recurring job scheduler	

Created	6 minutes ago
---------	---------------

The next image image shows the process graph:



Here the recurring Job:

You are using an unsupported command-line flag: `--disable-web-security`. Stability and security will suffer.

Hangfire Dashboard Jobs (0) Retries (0) Recurring Jobs (1) Servers (1) [Back to site](#)

Recurring Jobs

[Trigger now](#) [Delete](#)

Items per page: 10 20 50 100 500 1,000 5,000

<input type="checkbox"/>	Id	Cron	Time zone	Job	Next execution	Last execution	Created
<input type="checkbox"/>	BookIndexerJob	<code>0 0 * * *</code>	UTC	BookIndexer.Job	in 15 hours	5 minutes ago	2 days ago

Total items: 1

Hangfire 1.8.12 SQL Server: DEV01-WIN-SERV-2022@Libera HangfireServer 1.8.12 Storage Time: 12/5/2025 12:30:13 AM Application Time: 12/5/2025 12:30:13 AM Generated: 203 49ms

Here the Hangfire service definition for the Job:

```
var hangfireConnectionString = builder.Configuration.GetConnectionString("DefaultConnection");
builder.Services.AddHangfireInMemory();
SetCompatibilityLevel(Configuration => configuration
    .UseSimpleAssemblyNameTypeSerializer()
    .UseRecommendedSerializerSettings()
    .UseSqlServerStorage(
        hangfireConnectionString,
        new SqlServerStorageOptions
        {
            CommandBatchMaxTimeout = TimeSpan.FromMinutes(5),
            SlidingInvisibilityTimeout = TimeSpan.FromMinutes(5),
            QueuePollInterval = TimeSpan.Zero,
            UseRecommendedSetting = true,
            DisableGlobalLocks = true
        }
    ));
builder.Services.AddHangfireServer();
builder.Services.AddCors();
builder.Services.AddDistributedRedisExplorer();
builder.Services.AddSwaggerGen();
//Builder.Services.AddTransient<ITestRecurringJob, TestRecurringJob>();
//Builder.Services.AddScoped<ITestJob, TestJob>();
builder.Services.AddTransient<IBookIndexerJob, BookIndexerJob>();

var app = builder.Build();

// Configure the HTTP request pipeline.
if (app.Environment.IsDevelopment())
{
    app.UseSwagger();
    app.UseSwaggerUI();
}

app.UseHttpsRedirection();
app.UseControllers();

// Change 'Back to site' link URL
var options = new DashboardOptions { AppPath = "http://localhost:8200/" };
// If 'Back to site' link working for subfolder applications
//var options = new DashboardOptions { AppPath = VirtualPathUtility.ToAbsolute("~/") };

app.UseHangfireDashboard("hangfire", options);
RecurringJob.AddOrUpdate("BookIndexerJob", () => app.Services.GetRequiredService<IBookIndexerJob>().RunAsync(), Cron.Daily);

ann.Bin();

```

100% 0 issues found

Ready

File Edit View Project Build Debug Test Analyze Tools Extensions Window Help | Search | LibraSearch-v1.x

Solution Explorer

LibraSearch-HangfireServer (7 of 7 projects)

- Scripts
- LibraSearch-HangfireServer
- LibraSearch-ConsoleApp
- LibraSearch-Indexer
- LibraSearch-Model
- LibraSearch-WebAPI
- LibraSearch-Angular
- LibraSearch-SQL

Properties

Controllers

- BookIndexerController.cs
- TestJobController.cs
- TestRecurringJobController.cs
- TestRedisController.cs
- TestSqlController.cs

Pages

- Index.cshtml

Views

- BookIndexer
- TestJob
- TestRecurringJob

LibraSearch-Indexer

- Dependencies
- Extractor
- Program.cs
- Index.cs

LibraSearch-Model

- BookIndexerMigr.cs

LibraSearch-WebAPI

- Program.cs
- Index.cs

LibraSearch-Angular

- index.html

LibraSearch-SQL

GitHub Copilot ADMIN

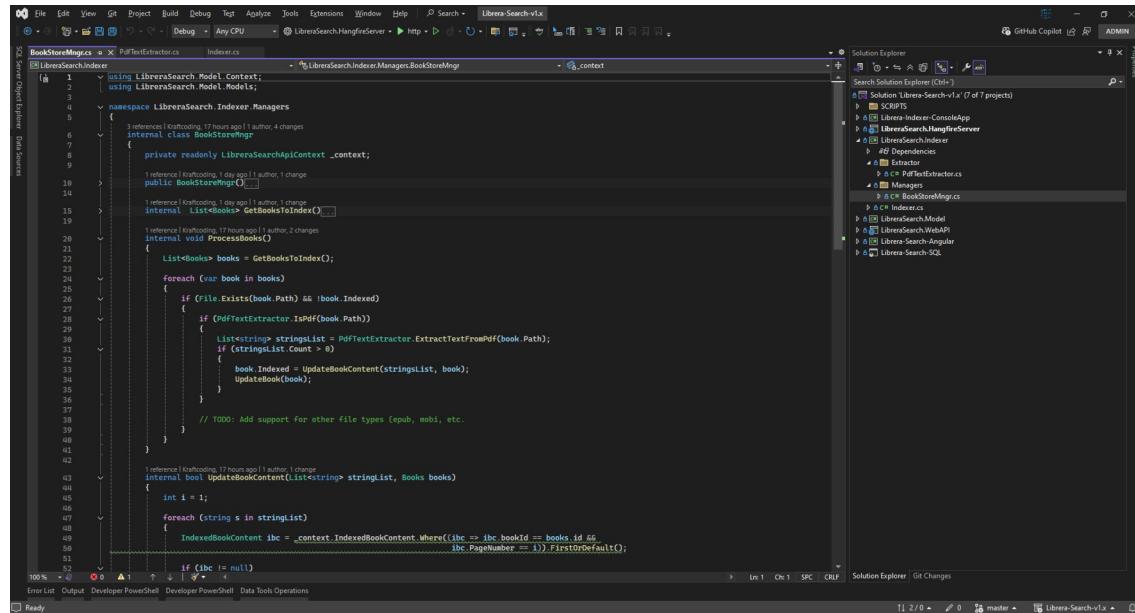
Ready

Output Developer PowerShell Developer PowerShell Data Tools Operations

11 2 / 0 0 0 master LibraSearch-v1.x

RQ-02 | Indexer

The indexer is implemented by a library type project on which we have the PDF extractor and the c# method that the Hangfire service will call by a recurring job definition.



```
1  using LibreraSearch.Model.Context;
2  using LibreraSearch.Model.Models;
3
4  namespace LibreraSearch.Indexer.Managers
5  {
6      internal class BookstoreMngr
7      {
8          private readonly LibreraSearchDbContext _context;
9
10         public BookstoreMngr()
11         {
12             _context = new LibreraSearchDbContext();
13         }
14
15         internal void GetBooksToIndex()
16         {
17             foreach (var book in books)
18             {
19                 if (File.Exists(book.Path) && !book.Indexed)
20                 {
21                     if (PdfTextExtractor.IsPdf(book.Path))
22                     {
23                         List<string> stringList = PdfTextExtractor.ExtractTextFromPdf(book.Path);
24                         if (stringList.Count > 0)
25                         {
26                             book.Indexed = UpdateBookContent(stringList, book);
27                         }
28                     }
29                 }
30             }
31         }
32
33         internal void ProcessBooks()
34         {
35             List<Book> books = GetBooksToIndex();
36
37             foreach (var book in books)
38             {
39                 if (File.Exists(book.Path) && !book.Indexed)
40                 {
41                     if (PdfTextExtractor.IsPdf(book.Path))
42                     {
43                         List<string> stringList = PdfTextExtractor.ExtractTextFromPdf(book.Path);
44                         if (stringList.Count > 0)
45                         {
46                             book.Indexed = UpdateBookContent(stringList, book);
47                         }
48                     }
49                 }
50             }
51         }
52
53         internal void UpdateBookContent(List<string> stringList, Books books)
54         {
55             int i = 1;
56
57             foreach (string s in stringList)
58             {
59                 IndexedBookContent ibc = _context.IndexedBookContent.Where(ibc => ibc.bookId == books.id &&
60                                         ibc.PageNumber == i).FirstOrDefault();
61
62                 if (ibc != null)
63                 {
64                     ibc.Content += s;
65                     ibc.PageNumber++;
66                 }
67             }
68         }
69     }
70 }
```

The indexer extraccts content form PDFs and stores it in DB.

RQ-03 | Model Layer

The model layer is managed by EF Core and LinQ is used to launch the queries.

RQ-04 | Web API

The Web API permits to the front-end application to execute CRUD operations my RESTful end-points. Those end-points are securitized by policy-based authorization.

Here some images.

Web API Auth. service setup:

```
1  using LibreraSearch.Models.Context;
2  using Microsoft.AspNetCore.Authentication.JwtBearer;
3  using Microsoft.EntityFrameworkCore;
4  using Microsoft.IdentityModel.Tokens;
5  using System.Text;
6
7  var builder = WebApplication.CreateBuilder(args);
8
9  // Add services to the container.
10
11 builder.Services.AddControllers();
12 // Learn more about configuring Swagger/OpenAPI at https://aka.ms/aspnetcore/swashbuckle
13 builder.Services.AddEndpointsApiExplorer();
14 builder.Services.AddSwaggerGen();
15
16 // DB Context
17 builder.Services.AddDbContext<LibreraSearchDbContext>(o => o.UseSqlServer(builder.Configuration.GetConnectionString("LibreraSearchDBConnString")));
18
19 // This code ensures that only users who possess valid JWT tokens issued by "YourIssuerName" & containing the "UserId" claim can access restricted s
20
21 builder.Services.AddAuthentication(JwtBearerDefaults.AuthenticationScheme)
22     .AddJwtBearer(options =>
23     {
24         options.TokenValidationParameters = new TokenValidationParameters
25         {
26             ValidateIssuer = true,
27             ValidateIssuerName = "YourIssuerName",
28             ValidateAudience = true,
29             ValidateLifetime = true,
30             IssuerSigningKey = SymmetricSecurityKey(Encoding.UTF8.GetBytes(builder.Configuration["Tokens:Key"]))
31         };
32     });
33
34
35 builder.Services.AddAuthorization(
36     options => options.AddPolicy("Authenticated",
37         policy => policy.RequireClaim("Email")));
38
39
40 var app = builder.Build();
41
42 // Configure the HTTP request pipeline.
43 if (app.Environment.IsDevelopment())
44 {
45     app.UseSwagger();
46     app.UseSwaggerUI();
47 }
48
49 app.UseHttpsRedirection();
50
51 app.UseAuthorization();
52
53 Ready.
```

Web API Tokens configuration:

```
Schema: https://www.schemastore.org/appsettings.json
1  {
2      "AllowedHosts": "*",
3      "ConnectionStrings": {
4          "LibreraSearchDBConnString": "Data Source=DEV01-WIN-SERV-2022"
5      },
6      "Logging": {
7          "LogLevel": {
8              "Default": "Information",
9              "Microsoft.AspNetCore": "Warning"
10         }
11     },
12     "Tokens": {
13         "Key": "ThisIsTheMostSecretStringEver123456789%",
14         "Issuer": "localhost",
15         "Audience": "http://localhost:4200/"
16     }
17 }
```

Web API securitized end-points:

The screenshot shows the Visual Studio IDE interface with the following details:

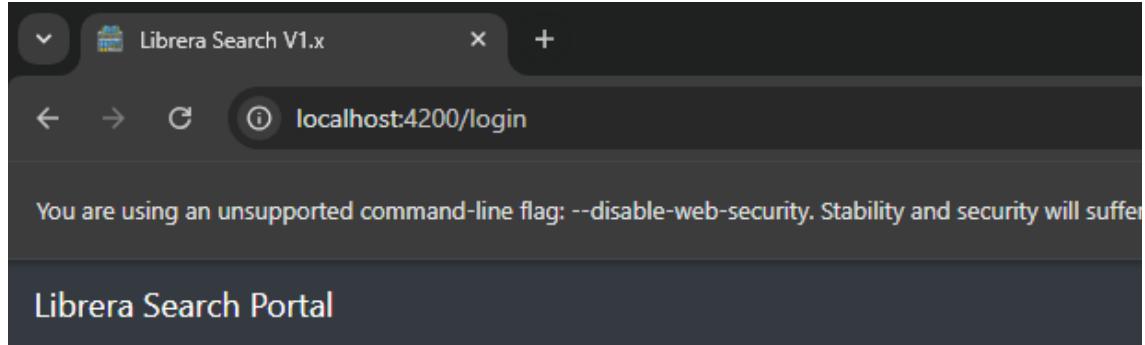
- Solution Explorer:** Shows the "Search Solution 'LibraSearch-v1.x'" which contains 7 projects:
 - Libra-Indexer-ConsoleApp
 - LibraSearch-ConfigServer
 - LibraSearch-Model
 - LibraSearch-WebAPI
 - Connected Services
 - Properties
 - LibraSearchController.cs
- Code Editor:** The file "LibraSearchController.cs" is open in the main editor window. It contains C# code for an ASP.NET Core Web API controller. The code includes imports for System, Microsoft, and LibraSearch namespaces, and defines a controller class "LibraSearchController" that inherits from "ControllerBase". It includes methods for GET and POST requests to manage books in a database context.
- Status Bar:** Shows the status bar with "100% 0" and "No issues found".
- Bottom Navigation:** Includes tabs for "Readme", "Output", "Developer PowerShell", "Data Tools Operations", "Solution Explorer", and "Git Changes".

RQ-05 | Front end App

The front end app is built in Angular 16.

Here some images.

Login form:



Login:

email	password	Submit
-------	----------	--------

Search by title list and details:

The screenshot shows a web-based application titled "Librera Search V1.x" running on "localhost:4200/LibreraSearch". The interface includes a top navigation bar with links for "Librera Search Portal", "Librera Search", "Add", "Text Content Search", "Hangfire Dashboard", and "Log out". Below the navigation is a search bar with a placeholder "Search by title" and a "Search" button. The main content area is divided into two sections: "Librera Search List" on the left and "Librera Search" details on the right.

Librera Search List

Kali inside Proxmox (Guest VM)

Linux Encrypted Filesystem with dm-crypt

Remove All

Librera Search

Modified: 2025-12-05T00:00:00
Title: Kali inside Proxmox (Guest VM)
Authors: Kali
Series: 1
Ids: isbn:0001
Published: Nov 10, 2025
Publisher: Linux
Languages: Eng
Tags: cybersecurity
Formats: pdf
Path: E:\TEST\PDF\Kali inside Proxmox (Guest VM).pdf
Indexed: false

Edit

Add form:

Librera Search V1.x

localhost:4200/add

Librera Search Portal Librera Search Add Text Content Search Hangfire Dashboard Log out

Modified

Title

Authors

Series 0

Ids

Published

Publisher

Languages

Tags

Formats

Path

Indexed

Submit

Search by content page details:

The screenshot shows a web browser window titled "Librera Search V1.x" with the URL "localhost:4200/LibreraTextSearch". The page header includes "Librera Search Portal", "Librera Search", "Add", "Text Content Search", "Hangfire Dashboard", and a "Log out" button. Below the header is a search bar with the word "store" and a "Search by Text Content" button. The main content area is titled "Librera Text Search List" and displays a single search result. The result includes the following fields and their values:

- Title:** Linux Encrypted Filesystem with dm-crypt
- Authors:** Joe Doe
- PageNumber:** 1

The detailed description of the result is as follows:

Linux Encrypted Filesystem with dm-crypt An encrypted filesystem will protect against bare-metal attacks against a hard drive. Anyone getting their hands on the drive would have to use brute force to guess the encryption key, a substantial hindrance to getting at your data. Windows and Mac OS X each provide its own standard cryptofs tools while Linux, of course, provides many tools to accomplish the task. The tool of choice these days, it seems, is dm-crypt . Invoked with the userspace cryptsetup utility, dm-crypt provides a fairly clean and easy-to-use cryptofs tool for Linux. Additionally, CentOS 5 includes an improved version of dm-crypt that supports LUKS . LUKS is an upcoming standard for an on-disk representation of information about encrypted volumes. Meta- data about encrypted data is stored in the partition header, and allows for compatibility between different systems and support for multiple user passwords. Besides that, GNOME and HAL have support for handling LUKS volumes, and can automatically prompt for a password if a removable medium with a LUKS volume is attached. If you do not require compatibility with older CentOS versions or systems that do not support LUKS, it is advised to use the LUKS scheme. The commands for setting up encrypted LUKS volumes are also described in the examples in this article. Here are Scripts to automate creation, un-mounting, and remounting of LUKS encrypted filesystems following the method described below.

1. Required Packages Before getting started, make sure all the requisite packages are installed:

- cryptsetup (cryptsetup-luks for CentOS-5)
- device-mapper
- util-linux

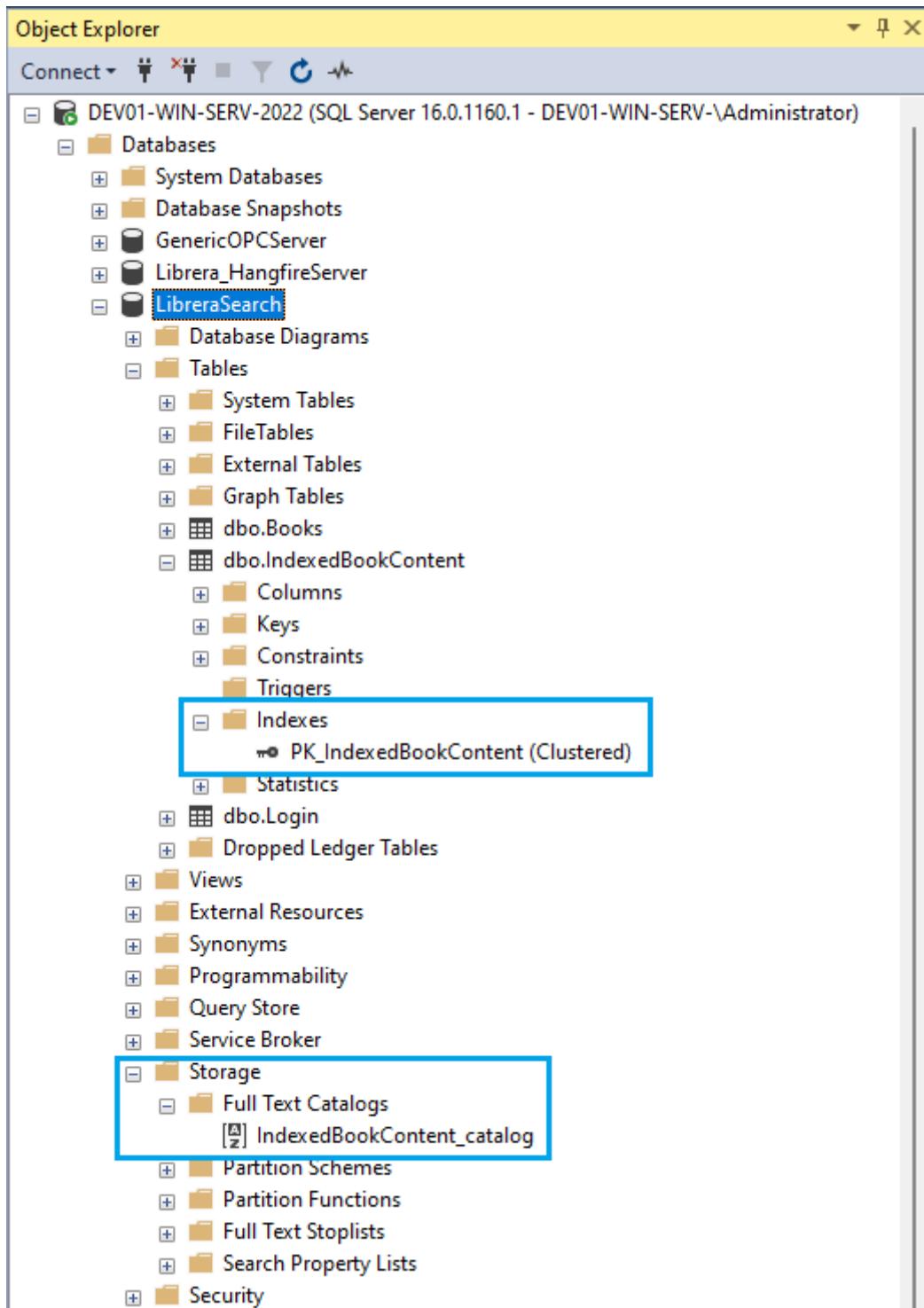
It's likely, however, that they're already present on your system, unless you performed a very minimal installation.

2. Initial FS Creation I typically encrypt files, not whole partitions, so I combine dm-crypt with the losetup loopback device maintenance tool. In the bare language of the Unix shell, here are the steps to create and mount an encrypted filesystem.

```
# Create an empty file sized to suit your needs. The one created # in this example will be a sparse file of 8GB, meaning that no # real blocks are written. Since we will force block allocation # lateron, it would not make much sense to do this now, since # the blocks will be rewritten anyway.
dd of=/path/to/secretfs bs=1G count=0 seek=8
# Lock down normal access to the file
chmod 600 /path/to/secretfs
# Associate a loopback device with the file
losetup /dev/loop0 /path/to/secretfs
```

RQ-06 | Database

The database is hosted by M. SQL Server 2022, all the schemas and T-SQL logic is in the project. Full-Text Search feature is needed, as we see in the next image:



Deployment

To do.

Software life cycle management

To do.

Annexes

A console App is added to test the indexer.

A script is added to launch Chrome avoiding CORS