







Book Data Pipeline

The Book Data Pipeline is a collaborative project developed by students at Northeastern University, specifically designed to enhance and streamline [ReMo](#)'s book metadata management system. It is a modern web application for managing and transforming book metadata, built with Next.js and MongoDB.

✨ Features

-  Support for multiple file formats (XML)
-  Batch processing of book metadata
-  Interactive book editing interface
-  Dark mode support
-  MongoDB integration for data persistence
-  Responsive design for all devices

✨ Screenshots

Book Data Pipeline

A modern solution for managing and transforming book metadata. Upload XML or Excel files, edit book information, and maintain a clean, structured database of your literary collection.

- XML Support
- Excel Support
- Batch Processing

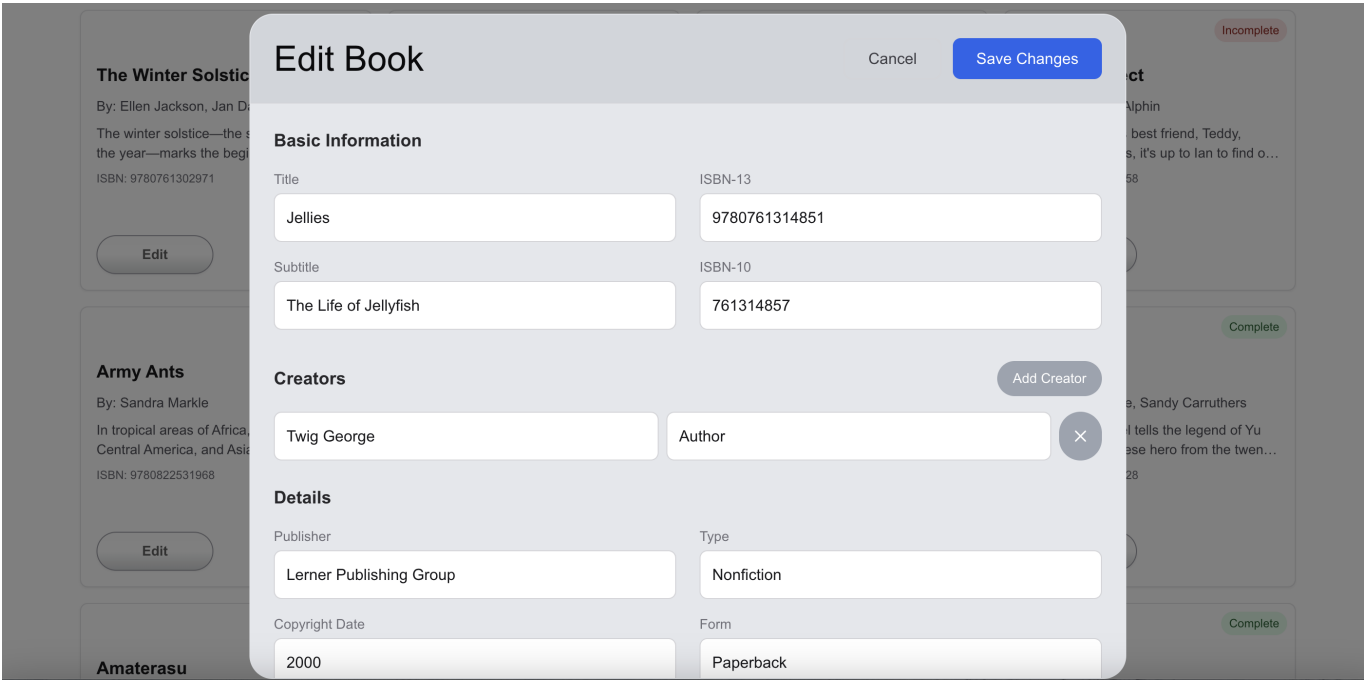
Database Status: Connected



Upload Book Data File (XML/XLSX)

Parse and Upload

<div>Incomplete</div> <div>The Winter Solstice</div> <div>By: Ellen Jackson, Jan Davey Ellis</div> <div>The winter solstice—the shortest day of the year—marks the beginning of the...</div> <div>ISBN: 9780761302971</div> <div>Edit</div>	<div>Complete</div> <div>Jellies</div> <div>By: Twig George</div> <div>"Exceptionally handsome photographs light up the pages of this well-designed...</div> <div>ISBN: 9780761314851</div> <div>Edit</div>	<div>Incomplete</div> <div>Africa Is Not a Country</div> <div>By: Margy Burns Knight, Mark Melnicov...</div> <div>Africa is not a country. From the tiny island nations of Comoros, Seychelles, and Sã...</div> <div>ISBN: 9780761316473</div> <div>Edit</div>	<div>Incomplete</div> <div>Picture Perfect</div> <div>By: Elaine Marie Alphin</div> <div>When Ian Slater's best friend, Teddy, suddenly vanishes, it's up to Ian to find o...</div> <div>ISBN: 9780822505358</div> <div>Edit</div>
<div>Incomplete</div> <div>Army Ants</div> <div>By: Sandra Markle</div> <div>In tropical areas of Africa, South America, Central America, and Asia, army ants ar...</div> <div>ISBN: 9780822531968</div> <div>Edit</div>	<div>Complete</div> <div>Captured by Pirates</div> <div>By: Justine Fontes, Ron Fontes, David ...</div> <div>Match your wits against a band of pirates on the high seas and try to defeat them...</div> <div>ISBN: 9780822562023</div> <div>Edit</div>	<div>Complete</div> <div>Hercules</div> <div>By: Paul D. Storrie, Steve Kurth</div> <div>Famous for his superhuman strength, Hercules is the most popular hero in...</div> <div>ISBN: 9780822564850</div> <div>Edit</div>	<div>Complete</div> <div>Yu the Great</div> <div>By: Paul D. Storrie, Sandy Carruthers</div> <div>This graphic novel tells the legend of Yu the Great, a Chinese hero from the twen...</div> <div>ISBN: 9780822565628</div> <div>Edit</div>



Getting Started

Prerequisites

- Node.js (v14 or higher)
- MongoDB instance (local or remote)
- npm or yarn package manager

Installation & Setup

1. First, set up the backend:

```
cd express
npm install
npm run dev
```

2. Then, in a new terminal(project root directory not in 'express' directory), set up the frontend:

```
# From the root directory
npm install
npm run dev
```

The application will be running at:

- Backend: <http://localhost:3000>
- Frontend: <http://localhost:3001>

Database Schema

The application uses MongoDB with the following book schema:

```

{
  title: {
    main: {
      type: String,
      trim: true,
      index: true      // Indexed for faster search
    },
    subtitle: {
      type: String,
      trim: true,
      default: ''
    }
  },
  creators: [{          // Authors, editors, etc.
    name: {
      type: String,
      trim: true
    },
    role: {
      type: String,
      trim: true
    }
  }],
  copyright_date: {     // Publication year
    type: Number,
    index: true,        // Indexed for faster search
    default: 0
  },
  summary: {            // Book description
    type: String,
    trim: true,
    default: ''
  },
  series: {
    name: {
      type: String,
      trim: true,
      default: ''
    },
    position: {         // Flexible type for different series numbering
      type: Mixed,
formats
      trim: true,
      default: ''
    }
  },
  genre: {
    main: {
      type: String,
      trim: true,
      index: true,      // Indexed for faster genre-based searches
      default: ''
    }
  }
}

```

```
    },
    subgenres: [{
      type: String,
      trim: true,
      default: ''
    }]
  },
  form: {                                // Book format
    type: String,
    trim: true,
    index: true,                          // Indexed for format-based filtering
    default: ''
  },
  pages: {
    type: Number,
    min: 0,                               // Cannot have negative pages
    default: 0
  },
  isbn: {
    isbn13: {
      type: String,
      sparse: true,                       // Sparse indexing for optional fields
      trim: true,
      default: ''
    },
    isbn10: {
      type: String,
      sparse: true,                       // Sparse indexing for optional fields
      trim: true,
      default: ''
    }
  }
}
```

The schema includes several indexed fields for optimized querying and search operations. All string fields are trimmed automatically to remove leading and trailing whitespace. Default values are provided for optional fields.

API Endpoints

- **GET** `/api/books/all` - Get paginated list of books
- **PUT** `/api/books/:id` - Update a book
- **DELETE** `/api/books/:id` - Delete a book
- **POST** `/api/books/add` - Add new books

Environment Variables

We suggest using the existing mongo in codebase for now, it will last till the end of 2024. Otherwise you need to match the database and collection name as well: [test] [books]

Create a `.env` file in the express directory with:

```
MONGODB_URI=your_mongodb_connection_string
```

Technologies Used

- Frontend: Next.js, React, Axios, Tailwind CSS
- Backend: Express.js, MongoDB, Mongoose
- Development: Node.js



Contributing

1. Fork the repository
2. Create your feature branch (`git checkout -b feature/AmazingFeature`)
3. Commit your changes (`git commit -m 'Add some AmazingFeature'`)
4. Push to the branch (`git push origin feature/AmazingFeature`)
5. Open a Pull Request



License

This project is licensed under the MIT License - see the [LICENSE](#) file for details.



Support

For support, please open an issue in the GitHub repository or contact the maintainers.