**DOI to Citation Converter with Excel Export**

**Overview**

The DOI to Citation Converter is an R Shiny web application that automatically converts Digital Object Identifiers (DOIs) into formatted citations with comprehensive bibliographic metadata. The application provides dual export capabilities in both RIS (Research Information Systems) format for citation managers and Excel format for data analysis and manipulation.

**Key Features**

**Core Functionality**

* **Batch DOI Processing**: Process multiple DOIs simultaneously (one per line input)
* **Dual Format Export**: Generate both RIS and Excel files from the same data
* **Abstract Integration**: Optional inclusion of abstracts when available
* **Real-time Preview**: Interactive data table showing processed citations
* **Progress Tracking**: Visual progress indicators and detailed status reporting
* **Error Handling**: Graceful handling of invalid DOIs or network failures

**Data Fields Extracted**

The application extracts comprehensive bibliographic metadata including:

| **Field** | **Description** | **Source** |
| --- | --- | --- |
| DOI | Digital Object Identifier | User input |
| Title | Publication title | CrossRef API |
| Authors | Author names (semicolon-separated) | CrossRef API |
| Journal | Journal/publication name | CrossRef API |
| Year | Publication year | CrossRef API |
| Volume | Journal volume number | CrossRef API |
| Issue | Journal issue number | CrossRef API |
| Pages | Page range (e.g., "123-135") | CrossRef API |
| Publisher | Publishing organization | CrossRef API |
| URL | Clickable DOI URL | Generated |
| Type | Publication type | CrossRef API |
| ISSN | International Standard Serial Number | CrossRef API |
| ISBN | International Standard Book Number | CrossRef API |
| Language | Publication language | CrossRef API |
| Abstract | Full abstract text (optional) | CrossRef API |
| Keywords | Subject keywords/terms | CrossRef API |
| Published\_Print | Print publication date | CrossRef API |
| Status | Processing status | Generated |

**Technical Architecture**

**Dependencies**

r

library(shiny) *# Web application framework*

library(rcrossref) *# CrossRef API integration*

library(openxlsx) *# Excel file generation*

library(dplyr) *# Data manipulation*

library(DT) *# Interactive data tables*

**Core Components**

**1. RIS Parser Function (parse\_ris\_to\_dataframe)**

A custom function that converts RIS format citations into structured data frames:

**Input**: RIS text string and DOI **Output**: Structured data frame with 19 bibliographic fields **Key Features**:

* Handles all standard RIS tags (TI, AU, JO, PY, etc.)
* Manages multiple authors and keywords
* Extracts year from various date formats
* Provides fallback values for missing data

**2. User Interface (UI)**

* **Input Panel**: DOI text area, abstract inclusion checkbox, submit button
* **Download Panel**: Separate buttons for RIS and Excel downloads
* **Main Panel**: Status reports, detailed logs, and interactive preview table

**3. Server Logic**

* **Reactive Processing**: Efficient handling of user inputs and data updates
* **Progress Monitoring**: Real-time feedback during batch processing
* **Data Storage**: Reactive values for citations and structured data
* **Error Management**: Comprehensive error handling and user notification

**Data Processing Workflow**

User Input (DOIs)

↓

CrossRef API Query (RIS format)

↓

RIS Parsing → Structured Data

↓

Optional Abstract Enhancement

↓

Data Validation & Storage

↓

Export Generation (RIS/Excel)

**Installation and Setup**

**Prerequisites**

* R (version 4.0 or higher)
* RStudio (recommended)
* Internet connection for CrossRef API access

**Required Packages Installation**

r

install.packages(c(

"shiny",

"rcrossref",

"openxlsx",

"dplyr",

"DT"

))

**Running the Application**

r

*# Load the application code*

source("doi\_converter\_app.R")

*# Or run directly if saved as app.R*

shiny::runApp()

**Usage Guide**

**Basic Operation**

1. **Enter DOIs**: Input one DOI per line in the text area
2. **Configure Options**: Check "Include Abstracts" if desired
3. **Process**: Click "Submit" to begin processing
4. **Monitor Progress**: Watch the progress bar and status messages
5. **Review Results**: Check the preview table for accuracy
6. **Download**: Choose RIS or Excel format for download

**Input Format Examples**

10.1038/nature12373

10.1126/science.1243259

10.1016/j.cell.2019.05.031

**Output Formats**

**RIS Format**

Standard citation manager format compatible with:

* Zotero
* Mendeley
* EndNote
* RefWorks

**Excel Format**

Structured spreadsheet with:

* Professional formatting
* Auto-sized columns
* Text wrapping for long fields
* Header styling
* Cell borders

**API Integration**

**CrossRef Integration**

The application leverages the CrossRef API through the rcrossref package:

* **Primary Endpoint**: Content negotiation for RIS format
* **Secondary Endpoint**: Works API for enhanced metadata
* **Rate Limiting**: Automatic handling of API constraints
* **Error Recovery**: Graceful degradation for failed requests

**Data Quality Considerations**

* **Coverage**: Limited to DOI-registered publications
* **Completeness**: Metadata availability varies by publisher
* **Accuracy**: Dependent on publisher-provided data
* **Timeliness**: Recent publications may have incomplete metadata

**Error Handling and Validation**

**Input Validation**

* Empty DOI detection
* Format verification
* Duplicate removal
* Whitespace trimming

**Processing Errors**

* Network connectivity issues
* Invalid DOI handling
* API rate limiting
* Malformed response data

**User Feedback**

* Real-time status updates
* Detailed error messages
* Success/failure counters
* Processing logs

**Performance Characteristics**

**Scalability**

* **Batch Size**: Tested with up to 100 DOIs
* **Processing Speed**: ~2-3 seconds per DOI
* **Memory Usage**: Efficient reactive programming
* **Network Dependency**: Requires stable internet connection

**Limitations**

* **API Dependencies**: Subject to CrossRef service availability
* **Rate Limits**: CrossRef API constraints may slow processing
* **Data Coverage**: Limited to DOI-registered publications
* **Browser Compatibility**: Requires modern web browser

**File Output Specifications**

**RIS Files**

* **Extension**: .ris
* **Encoding**: UTF-8
* **Structure**: Standard RIS format with enhanced fields
* **Compatibility**: Universal citation manager support

**Excel Files**

* **Extension**: .xlsx
* **Format**: Modern Excel format
* **Features**: Formatted headers, auto-sizing, text wrapping
* **Compatibility**: Excel 2007+, LibreOffice Calc, Google Sheets

**Future Enhancements**

**Potential Features**

* **Multi-database Integration**: PubMed, Scopus, Web of Science
* **Citation Network Analysis**: Reference and citation mapping
* **Duplicate Detection**: Automatic identification of duplicate entries
* **Batch Upload**: File-based DOI input
* **Export Templates**: Customizable output formats
* **User Authentication**: Personal libraries and history
* **API Keys**: User-provided API credentials for higher rate limits

**Technical Improvements**

* **Caching**: Local storage of processed citations
* **Parallel Processing**: Concurrent API requests
* **Database Backend**: Persistent storage capabilities
* **RESTful API**: Programmatic access to functionality

**Contributing and Development**

**Code Structure**

├── ui.R # User interface definition

├── server.R # Server logic

├── parse\_ris\_to\_dataframe.R # RIS parsing function

├── README.md # Documentation

└── tests/ # Unit tests (future)

**Development Guidelines**

* Follow R coding standards
* Implement comprehensive error handling
* Include inline documentation
* Test with diverse DOI types
* Maintain backward compatibility

**Citation and Attribution**

**Developed By**

**Dr. MIM. Riyath**  
Department of Accountancy and Finance  
Faculty of Management and Commerce  
South Eastern University of Sri Lanka  
Email: [riyath@seu.ac.lk](mailto:riyath@seu.ac.lk)

**License**

This software is provided under [specify license - e.g., MIT, GPL-3, etc.]

**Acknowledgments**

* CrossRef for providing comprehensive metadata API
* R Shiny team for the web application framework
* rcrossref package maintainers for API integration tools

**Support and Contact**

For technical support, bug reports, or feature requests:

* **Email**: [riyath@seu.ac.lk](mailto:riyath@seu.ac.lk)
* **Institution**: South Eastern University of Sri Lanka
* **Department**: Accountancy and Finance

**Version History**

**Version 1.0 (Current)**

* Initial release with core functionality
* RIS and Excel export capabilities
* Abstract integration
* Interactive preview table
* Comprehensive error handling