



## Sophisticated Online Tool to Sort Colon Classification Number

Sreekumar S<sup>a</sup>, Suresh Yadav<sup>b</sup> and Selva Kumar S<sup>c</sup>

<sup>a</sup> Assistant Librarian, Delhi University Library System, University of Delhi, Delhi, Pin 110007, India.

Email: sreekumar@drtc.isibang.ac.in

<sup>b</sup> Research Scholar, Department of Library and Information Science, Banasthali Vidyapith, Rajasthan, Pin 304022, India.

Email: sureshy@drtc.isibang.ac.in

<sup>c</sup> Assistant Librarian, Delhi University Library System, University of Delhi, Delhi, Pin 110007, India.

Email: skumar@duls.du.ac.in

*Received: 19 December 2025; Accepted: 19 February 2026*

Colon Classification (CC), developed by Dr. S. R. Ranganathan, was one of the underrated classification systems, which seems to be on the brink of death. It is still being adopted and followed by some of the significant academic and public libraries in India and abroad. Even though being positioned in the top classification list, the complexity of the Colon Classification's Mixed notation makes it difficult for the New Libraries to adopt the Colon Classification as their Classification system. Due to the complexity of the mixed notations, CC really needs a qualified, experienced Library patron to arrange the shelves. Even the Patrons found it challenging to locate the books compared to other classification systems. To address this issue, the paper introduces the "CC Sorter," which will provide an efficient mechanism and workflow for arranging books on shelves and may be considered a 21st-century revival of the Colon Classification. It overcomes errors in shelving and book identification, enhancing the overall experience for both library staff and patrons.

**Keywords:** Colon Classification, Library, Sorting Tool, Shelf Arrangement, User Navigation

### 1. Introduction

Colon Classification (CC)<sup>1</sup>, developed by Dr. S. R. Ranganathan, is one of the most flexible classification schemes used in libraries. It has several facets, which make the classification system more precise and adaptable. However, the mixed notations comprising different characters, such as Numbers, letters, Punctuation marks, and symbols, each with its own precedence, make it complicated to sort the collection manually<sup>2</sup>. This results in an incorrect arrangement of books and is time-consuming, violating the "APUPA (Alien-Penumbral-Umbral-Penumbral-Alien) Pattern" and "Five Laws of Library Science"<sup>3</sup>.

Even Computers fail to sort CC numbers because they follow lexical and numerical rules. As there is a lack of automated methodologies to interpret CC Numbers, librarians rely on labor-intensive manual sorting. Consider an example where you are asked to provide a list of books on the Mathematical topic "Riemann", with its respective CC number being "B37:1M5". However, when sorting in the computer system, it doesn't always provide the desired results. Even if it allows it, it won't be in the proper sequence.

This paper proposes a solution to an issue that has persisted for an extended period, which remains unaddressed. This tool can sort CC numbers in the exact filiation sequence according to the CC rule.

### 2. Literature Review

Colon Classification is one of the cornerstones of library organization, enabling the systematic arrangement and retrieval of documents. Over time, various Library Classification schemes, such as the Dewey Decimal Classification (DDC), Universal Decimal Classification (UDC), and Library of Congress Classification (LCC), have been developed to classify knowledge resources. These systems are enumerative or hierarchical, making them easy to sort in digital environments. In recent years, as the adoption of Integrated Library Systems (ILS) such as Koha, Evergreen, OpenBiblio, Invenio, and New GenLib has increased, the need for precise sorting and Shelf arrangement has become more vital across all classification schemes.

Among these classification schemes, the Colon Classification (CC) developed by S. R. Ranganathan stood out for its Analytico-synthetic method based on

the PMEST formula – Personality, Matter, Energy, Space, and Time. It was developed in the year 1933. CC has showcased a drastic shift from an enumerative classification scheme to a faceted approach, which can accommodate and express any complex relationships<sup>4</sup>. It contains several notations comprising of alphabets, numbers, punctuation marks, and special characters, which all provide a multidimensional subject relationship. CC being theoretically rich, the syntactic complexity poses difficulty in automation and digital implementation.

Several studies have contributed to the study of Colon Classification. M. P. Satija (1997, 2011, 2013)<sup>5,6,7</sup> provided extensive interpretations on Colon Classification, emphasizing its relevance and adaptation in the digital era. Some efforts have been made in the past by Panigrahi (2005)<sup>8</sup> to automate the Colon Classification. In recent years, Halder (2023)<sup>9</sup> has also made an effort to generate Colon Class Numbers for Indian Literary Works. However, most works focus on theoretical foundations, with limited attention to computation or sorting.

In summary, the literature review reveals that the Colon Classification has been extensively studied from both conceptual and historical perspectives; however, only a few works have addressed automating the Colon Classification. Among which, no effort has been taken to automate its sorting logic for the digital environment. Most of the Integrated Library Management Systems are designed for simpler schemes and do not support the Colon Classification scheme. This gap shows the need to develop a dedicated Colon Classification Sorter capable of sorting complex CC Syntax.

### 3. Problem Statement

Although the Colon Classification is widely used in Libraries, no tool is currently available to correctly sort CC numbers. For the Dewey Decimal Classification (DDC), a spreadsheet software can sort the DDC Numbers in Numerical order. Unlike other classification systems, the Colon Classification employs a complex mixed notation in classifying the universe of knowledge, making it difficult to sort by spreadsheet applications or other software. Even testing a few datasets with publicly available AI tools resulted in incorrect arrangements of the CC numbers. Existing tools, such as spreadsheet applications, AI tools, and library management systems, failed to sort CC numbers.

This violates the APUPA pattern, a fundamental principle of Colon Classification, resulting in incorrect shelf arrangements and subject closeness. Without a proper sorting algorithm, this problem will remain unaddressed.

#### 3.1 Implications of the Problem

- **Reduced Efficiency:** Manual sorting takes time, and finding the correct shelves or the location where it can be placed.
- **Inconsistent Shelf Arrangements:** In some cases, the patron/staff fails to recognize the location of the book due to complex mixed notations and places it in an incorrect place, which in turn leads to the missing of books and unreliable catalog/bay guides.
- **Limited Adoption:** Due to the complexity of mixed notations in CC, it discourages other libraries from adopting CC as their Classification system.

### 4. Methodology

The Sorting Methodology identifies all characters specified in the Colon Classification and their precedence according to the CC Filiatory sequence<sup>10</sup> (Fig. 1). CC has its own unique hierarchy, where every character carries a context-specific meaning and relationship.

#### i. Define Precedence Order:

Establish a priority order for characters: small alphabets > zero > inverted comma > dot > colon > semicolon > dash > numbers > main class digits > arrow marks > brackets

The arrays of the individual character sets are defined in their respective priority lists. For example, all the small letters from ‘a’ to ‘z’ are specified in the array set ‘small Alphabets’. Table 1 provides the hierarchical order of CC numbers and their respective values.

#### ii. Precedence Mapping

Each character specified in the array is assigned a rank in the precedence order. The rank assigned to the characters is used to compare the strings during the sorting process.

#### iii. Comparison Function

The comparison function compares the two CC numbers, string by string, using the logic:



Fig. 1 — Methodology of Colon Classification Sorter Tool

Table 1 — Precedence Order according to the Filiatory Sequence of Colon Classification

Priority	Digits	
1	Small alphabets	a – z
2	Arrow marks	← →
3	Zero	0
4	Inverted comma	'
5	Dot	.
6	Colon	:
7	Semi Colon	;
8	Comma	,
9	Dash	-
10	Numbers (1-9)	1 - 9
11	Capital Letters	A - Z
12	Brackets	()

```
function compareStrings(s1, s2):
    arr1 = split(s1)
    arr2 = split(s2)
    minLength = min(length(arr1), length(arr2))

    for i from 0 to minLength - 1:
        rank1 = precedenceMap[arr1[i]]
        rank2 = precedenceMap[arr2[i]]

        if rank1 != rank2:
            return rank1 - rank2

        if i < length(arr1) - 1 and arr1[i+1] in (smallAlphabets ∪ arrowMarks):
            if i >= length(arr2) - 1 or arr2[i+1] not in (smallAlphabets ∪ arrowMarks):
                return -1

        if i < length(arr2) - 1 and arr2[i+1] in (smallAlphabets ∪ arrowMarks):
            if i >= length(arr1) - 1 or arr1[i+1] not in (smallAlphabets ∪ arrowMarks):
                return 1

    return length(arr1) - length(arr2)

sortedArray = usort(inputArray, compareStrings)
```

Fig. 2 — Pseudocode of Colon Classification Sorter Tool

- Comparing the ranks of the corresponding characters of the two strings using precedence mapping as specified above.

- If the ranks are different, then the string that had the lower rank character will come first.

#### iv. Sorting Algorithm

The Colon Classification sorting algorithm performs the sorting after the strings have been compared through the comparison function. The pseudocode of the CC sorter is as show in the Fig. 2.

### 5. The Solution: An Online Sorting Tool

An Online Sorting Tool was developed to address the long-standing issue of sorting of Colon Classification Numbers. This tool sorts the CC numbers according to the Colon Classification filiatory sequence. It is made publicly available at <http://colonclassify.in>.

#### 5.1 Key Features

**i. User-Friendly Interface:** This tool is developed with a user-centric design and ease of access.



Fig. 3 — Workflow of Colon Classification Sorter Tool

#### ii. Automated CC Sorting:

- **Multiple Entries:** The User can input multiple Colon Numbers if they have any doubts or confusion to verify the Shelf arrangements. It is ideal for a small dataset.

- **Bulk Entries:** For larger sets of data, users can upload their Excel file (up to 200MB) containing the CC numbers. It allows users to select the column in which the CC sorting should be performed. This sorting not only sorts the individual column but also provides an expanded column range sorting, which in turn doesn't alter the sequence of the data uploaded.

Additionally, this tool provides sorting options for both ascending and descending order. Once the results are generated, they can be exported as a CSV file. This tool is hosted on the web server <http://colonclassify.in> to make it globally accessible.

This tool was designed with the ease of access in mind for users with varying technical backgrounds. As it is web-hosted, the tool requires no installation and can be accessed from anywhere, thereby enhancing its usability and accessibility.

#### 5.2 Technical Implementation

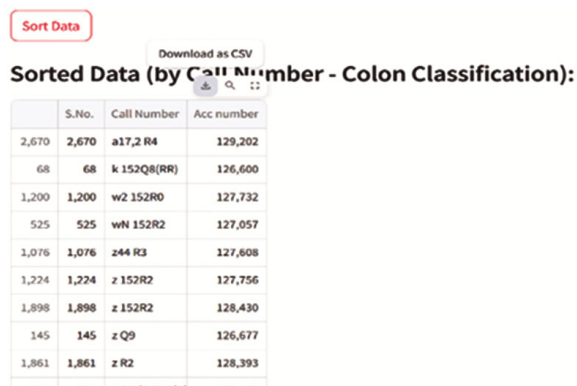
For implementing the Colon Classification Sorter, Python has been used for handling Programming Logic and functions. For data processing, we have used NumPy for numerical computations, such as arrays and matrices, and Pandas for filtering, grouping, aggregating, data cleaning and reading large datasets.

While Streamlit and CSS is used for displaying the Python scripts into an interactive web application. Workflow is shown in Fig. 3.

#### 5.3 Impact and Benefits

##### i. Enhanced Efficiency

Sorting of books will take less time than usual. The Library professionals, who were previously deployed to sort and classify numbers in order, can now focus on significant aspects of the library, leading to increased operational productivity.



S.No.	Call Number	Acc number
2,670	a17,2 R4	129,202
68	k 152Q8(RR)	126,600
1,200	w2 152R0	127,732
525	wN 152R2	127,057
1,076	z44 R3	127,608
1,224	z 152R2	127,756
1,898	z 152R2	128,430
145	z Q9	126,677
1,861	z R2	128,393

Fig. 4 — Final Results After Sorting the Colon Classification Numbers

## ii. Consistency and Accuracy

It provides accurate results, reducing the possibility of human errors and ensuring consistency in the classification/shelving order.

## iii. Wider Accessibility

The tool is available online, allowing users to access it from anywhere. The notational system, which was once a barrier to adoption due to its complex notation, is now made easier. It bridges the technical gap; the CC Sorter Tool promotes the continued use of Colon Classification and the relevance of Ranganathan's system.

## 6. Results

The CC Sorter tool has been tested with a dataset containing 2750 CC numbers and their respective accession numbers (Fig. 4). The tool successfully sorted all the CC numbers in the column, retaining their associated columns intact, i.e., Accession Number. The results were processed within a few seconds, demonstrating the tool's computational ability and robustness.

Furthermore, the tool was tested with a small dataset, which also yielded accurate results, proving the CC Sorter's reliability and scalability across varying input sizes.

## 7. Future directions

- **Integration with Library Management Systems:** In the future, the tool can be integrated with the Library Management Systems to provide a proper shelf list and enhance its utility.
- **User Navigation & Book Shelf Location Identification:** The CC Sorter Algorithm can be used in Libraries using Colon Classification for Book Location Identification.

- **Expansion of Features:** Additional features, such as automated generation of CC numbers, or recommendation of Classification Numbers from the expert groups.

## 8. Conclusion

The CC Sorter Tool represents a significant advancement in the field of library classification systems, particularly for libraries using Colon Classification. With the automation of the sorting process, the tool addresses critical challenges in shelf rectification, library organization, and user navigation. It can be further integrated into library management software, such as Koha, to enhance its utility even further, ensuring accurate and automated shelf lists.

As one of the world's most renowned classification systems, Colon Classification deserves continued attention and modernization. With millions of books classified under this system and being used globally, tools like CC Sorter play a pivotal role in preserving its relevance and efficiency in the digital age.

## 9. References

- 1 Ranganathan, S. R., Colon classification, 6<sup>th</sup> edn (Asia Publishing House; Bombay), 1963.
- 2 Boyce, B. R., & Stowe, M. (n.d.). A brief investigation of U.S. superintendent of documents filing practice. *Government Publications Review*, 18(4), 347–352. [https://doi.org/10.1016/0277-9390\(91\)90030-2](https://doi.org/10.1016/0277-9390(91)90030-2)
- 3 Ranganathan, S. R., The Five Laws of Library Science, (Madras Library Association; New Delhi), 1931.
- 4 Satija, M. (2017). Reviews of Concepts in Knowledge Organization 291 M. *Knowledge Organization*, 44(4), 291–307. <https://doi.org/10.5771/0943-7444-2017-4-291>
- 5 Satija, M. P. (1997). The Revision and Future of Colon Classification. *Knowledge Organization*, 24(1), 18–23.
- 6 Satija, M. P., A guide to the theory and practice of Colon Classification, (Ess Ess Publications, New Delhi), 2011.
- 7 Satija, M. P. (2013). Colon Classification: A Requiem. *DESIDOC Journal of Library & Information Technology*, 33(4), 265–276.
- 8 Panigrahi, P., & Prasad, A. R. D. (2005). Inference Engine for devices of Colon Classification in AI Based Automated Classification System. *Annals of Library and Information Studies*, 52(4), 124–134.
- 9 Halder, D., & Biswas, M. (2023). Machine-generated colon class numbers: Automatic classification of Indian literary works in the wikidata environment. *Journal of Information and Knowledge*, 143–149. <https://doi.org/10.17821/srels/2023/v60i3/171025>
- 10 Bianchini, C. (2011). Organizing knowledge with the filiatory sequence of the Colon Classification. *JLIS: Italian Journal of Library, Archives and Information Science = Rivista Italiana Di Biblioteconomia, Archivistica e Scienza Dell' informazione*: 2, 2, 2011, 2(2), 1–21. <https://doi.org/10.4403/JLIS.IT-4710>