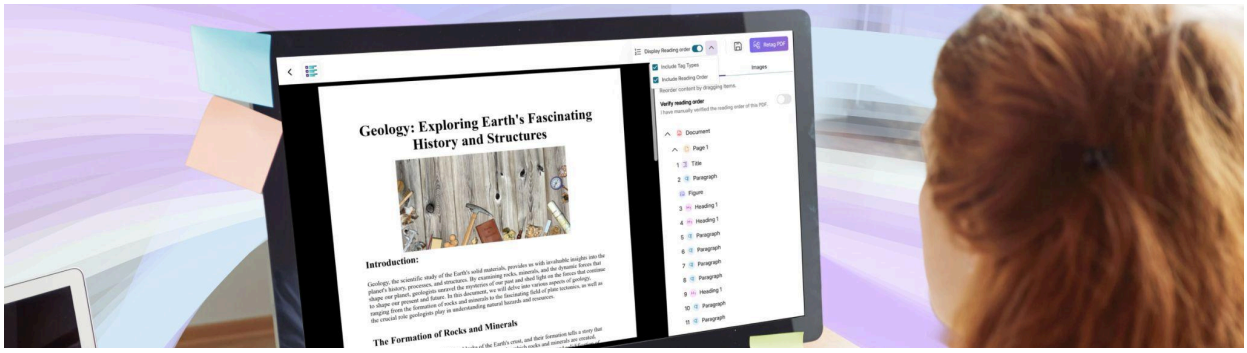
# Introduction




# Introduction

Accessible PDFs allow everyone, including individuals using screen readers or other assistive technologies, to access the same information, in the same logical order, as sighted readers. A PDF may look fine visually but still be completely inaccessible if it hasn't been properly structured.

Three main concepts form the foundation of PDF accessibility:

- Optical Character Recognition (OCR) has been applied
- Page elements are properly tagged
- Elements are assigned in the correct reading order

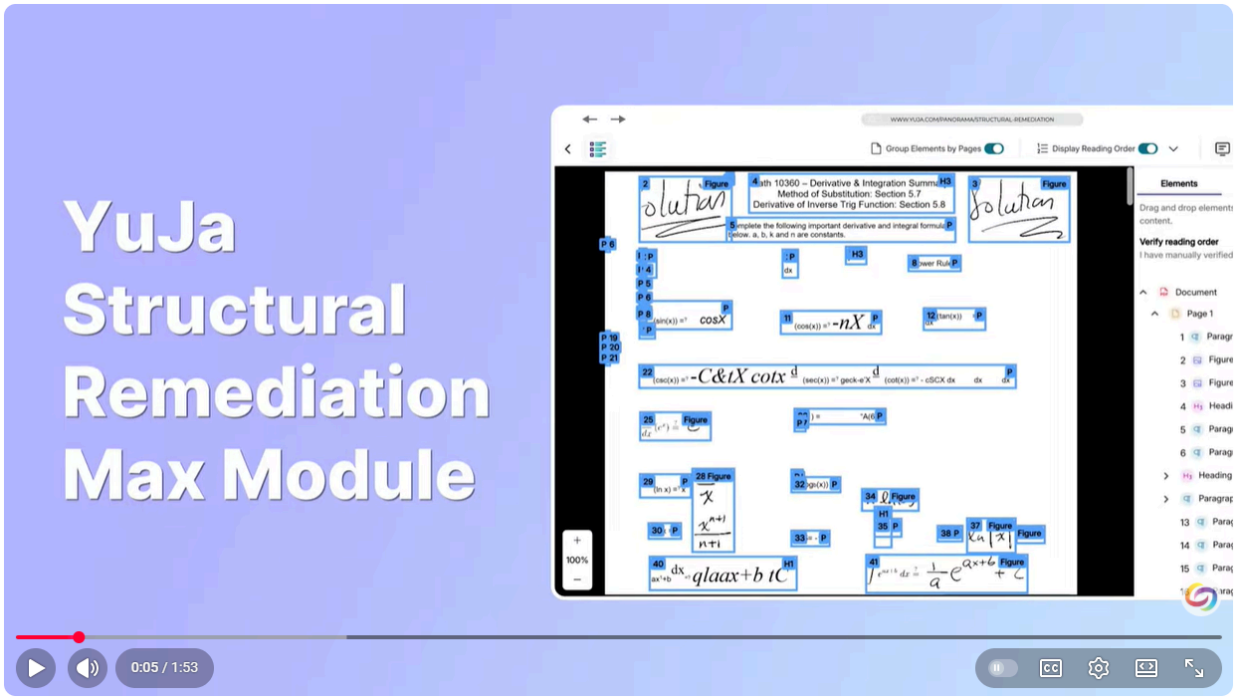
Depending on how complex the document is, some issues may be simple to fix, while others will require more time. PDFs with many page elements, low-quality scans, detailed graphics, or handwritten notations are more likely to generate a wide range of accessibility errors.

## YuJa Panorama Structural Remediation Max

Since PDF was originally created by Adobe, solutions for fixing existing accessibility errors have typically required a paid license for Adobe Acrobat Pro. Elgin Community College recently purchased an alternative solution that is now part of our existing suite of YuJa Panorama tools.

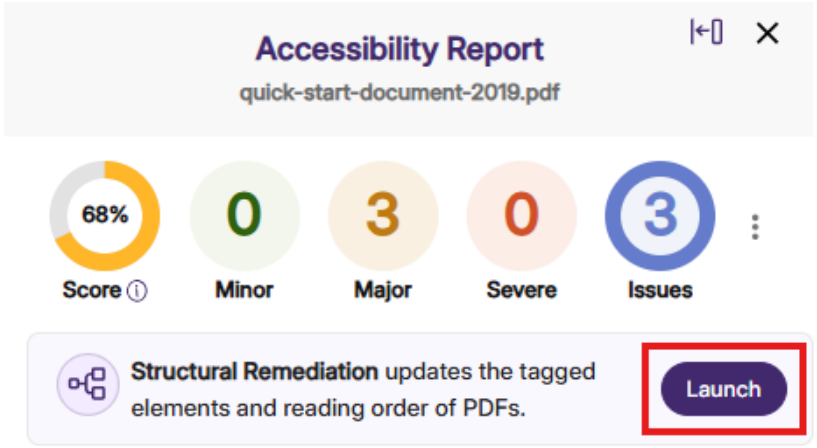
YuJa Panorama's Structural Remediation Max helps faculty make existing PDFs more accessible by adding the proper structure that screen readers rely on. It uses AI to recognize the layout of a document and scan and apply tags such as headings, paragraphs, lists, tables, and images. While the tagging isn't always perfect and may require review or adjustment, it provides the features needed to remediate scanned or untagged PDFs for much greater accessibility that have been previously unavailable.

**Selecting the image below will launch the Welcome to the YuJa Structural Remediation Max Module video:**

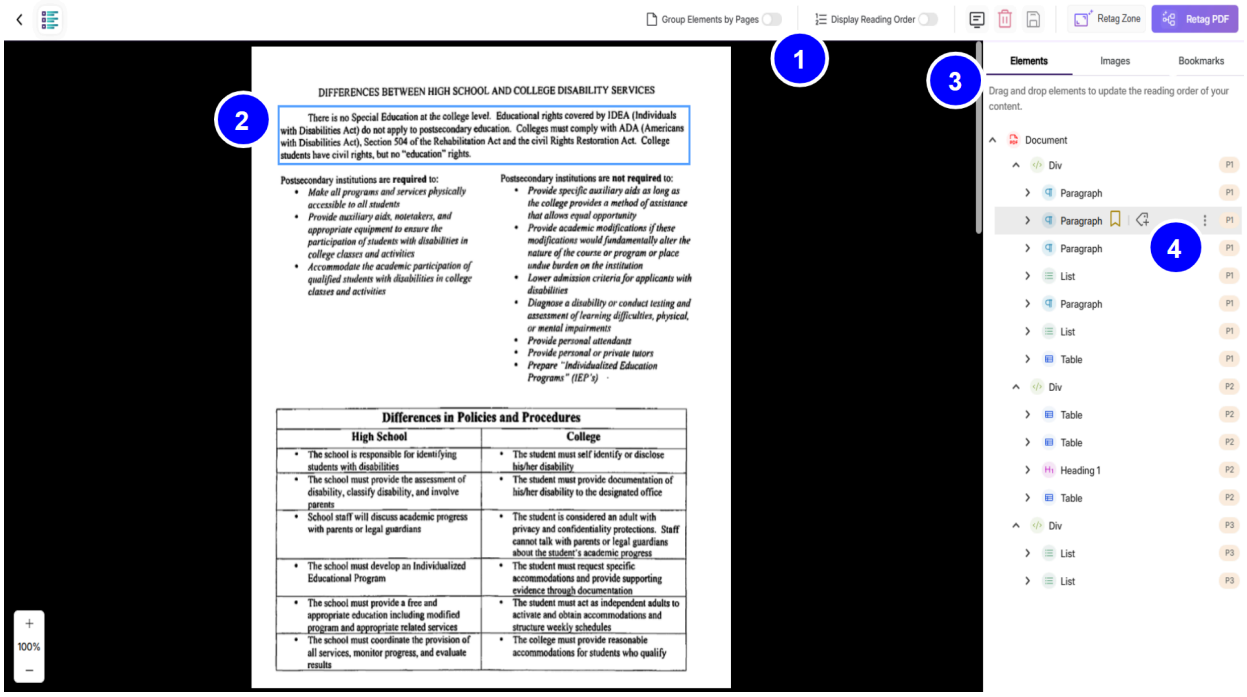



# Getting Around Structural Remediation Max

Within the YuJa Panorama Accessibility Report an option to Launch will appear. Selecting this will open the Structural Remediation dashboard.



Below is an example of the dashboard with some of the key features marked and described.




## 1. Group Elements by Page and Display Reading Order

- Group Elements by Page organizes the list of tagged elements one page at a time, rather than showing every element from the entire PDF in a long list. This allows users to focus on making fixes page-by-page, which can be helpful when working with longer documents.

- Display Reading Order will highlight all tagged elements in bounding boxes and display numbers in the top left corner of each box. These numbers correspond to the order of the elements in the reading order listed in the elements panel.

## 2. PDF Display

- Bounding boxes will appear around elements as they are selected in the elements panel, or all bounding boxes for all elements will show if the Display Reading Order toggle is on.

## 3. Elements Panel

- The main portion of the right sidebar displays the elements or tags present in the PDF.

- Hovering over the element will change the bounding box in the PDF display to indicate the location of the tag.
- Clicking the dropdown arrow can expand the view to show nested elements.

4. **Tags Order**
   - Clicking the three dots in a tag allows users to edit the tagged element type and also provides the ability to remove elements that are unnecessary for structure or content.
   - Hovering through the list of tagged elements or toggling on the Display Reading Order option and matching the assigned numbers can allow users to verify that content is structured in the appropriate reading order.

---

# Key Points About Structural Remediation Max

- Structural Remediation Max uses AI to apply Optical Character Recognition and detect structure and assign tags automatically. This process can significantly enhance the accessibility of the file, but it is not perfect.
- The AI does not train or learn from interactions.
- Just like any tool or software, there is a slight learning curve and using this tool gets easier with practice.
- The following workflow is recommended:
  1. Scan with OCR
  2. Review tagged structural elements
  3. Verify reading order
- A clean, well-structured source document is critical for the tool to work properly. If the initial scan is poor quality, consider the following:
  - Search the ECC library database for a cleaner, existing version of the document.
  - Consider creating a clean version with the same content in Microsoft Word or Google Docs and exporting as a PDF.

- AI-generated tags still need human review. Check the structure panel after applying tags as some tags may need to be:
    - **Changed** - a tagged element was assigned an incorrect element type
    - **Removed** - decorative lines or unnecessary elements incorrectly tagged
    - **Added** - a page element is not displaying an associated tag

- This tool may apply advanced or specialized tag types. While the system supports many PDF tag types, it is recommended to focus on six core structural elements.
    - **Paragraph** - body text
    - **Headings** - titles of sections and subsections for structure and navigation
    - **Figures** - images and diagrams
    - **Lists** - organized information
    - **Tables** - structured data
    - **Links** - clickable resources

- Contact Instructional Technology & Accessibility Coordinators for additional support remediating PDFs or using Structural Remediation Max
    - Ryan Karp 847-214-7101 rkarp@elgin.edu
    - Kyle Ahlmann 847-214-7476 kahlmann@elgin.edu

# Scans and OCR

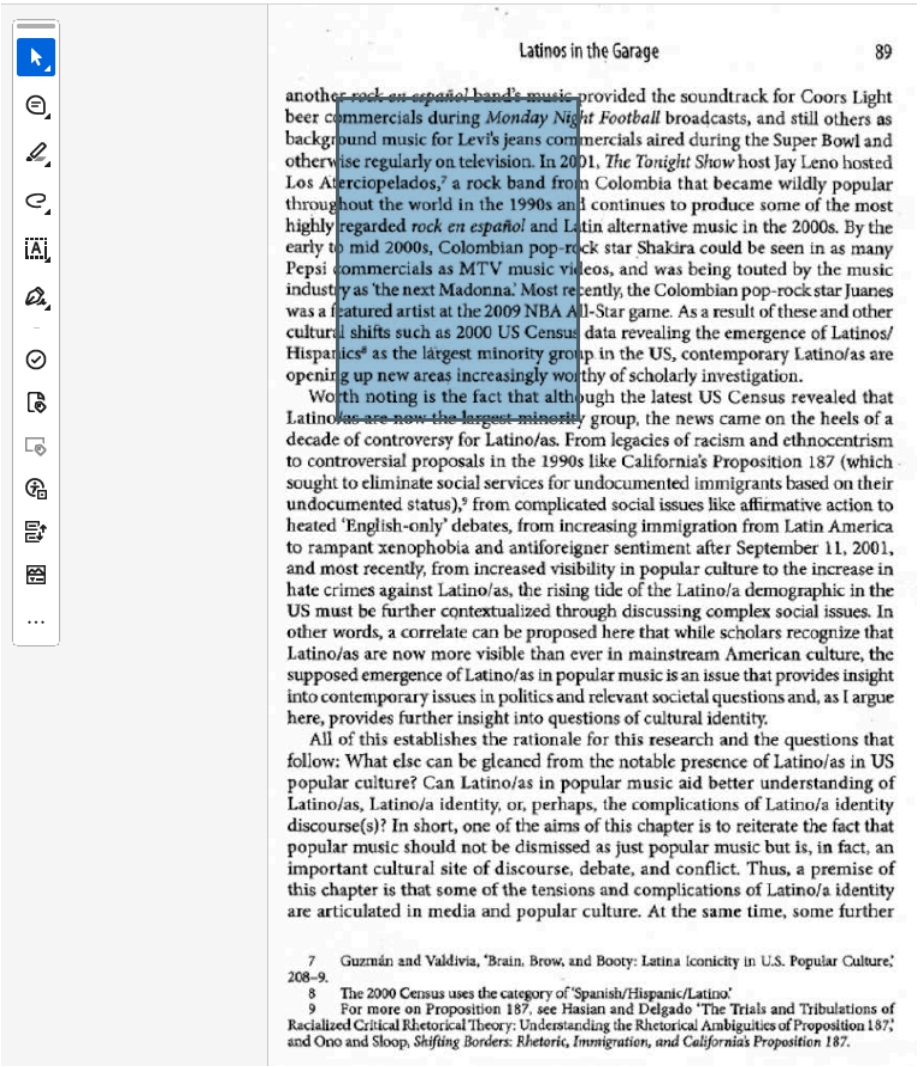

# Scans and Optical Character Recognition

## What “Scanned PDF” Means

A scanned PDF is essentially a digital image of a physical document. Instead of containing selectable, digital text, it contains a picture of the text.

Accessibility Implications for Scanned PDFs:

- **0% Accessible:** Assistive technologies cannot read text in scanned PDFs because the content is just an image.
- **Navigation is Broken:** Headings, lists, tables, and other structural elements cannot be detected.
- **Search and Copy Functions Fail:** Users cannot search for text, highlight it, or copy it for necessary purposes.



Latinos in the Garage 89

another rock en español band's music provided the soundtrack for Coors Light beer commercials during *Monday Night Football* broadcasts, and still others as background music for Levi's jeans commercials aired during the Super Bowl and otherwise regularly on television. In 2001, *The Tonight Show* host Jay Leno hosted Los Aterciopelados,[7] a rock band from Colombia that became wildly popular throughout the world in the 1990s and continues to produce some of the most highly regarded *rock en español* and Latin alternative music in the 2000s. By the early to mid 2000s, Colombian pop-rock star Shakira could be seen in as many Pepsi commercials as MTV music videos, and was being touted by the music industry as 'the next Madonna.' Most recently, the Colombian pop-rock star Juanes was a featured artist at the 2009 NBA All-Star game. As a result of these and other cultural shifts such as 2000 US Census data revealing the emergence of Latinos/Hispanics[8] as the largest minority group in the US, contemporary Latino/as are opening up new areas increasingly worthy of scholarly investigation.

Worth noting is the fact that although the latest US Census revealed that Latino/as are now the largest minority group, the news came on the heels of a decade of controversy for Latino/as. From legacies of racism and ethnocentrism to controversial proposals in the 1990s like California's Proposition 187 (which sought to eliminate social services for undocumented immigrants based on their undocumented status),[9] from complicated social issues like affirmative action to heated 'English-only' debates, from increasing immigration from Latin America to rampant xenophobia and antiforeigner sentiment after September 11, 2001, and most recently, from increased visibility in popular culture to the increase in hate crimes against Latino/as, the rising tide of the Latino/a demographic in the US must be further contextualized through discussing complex social issues. In other words, a correlate can be proposed here that while scholars recognize that Latino/as are now more visible than ever in mainstream American culture, the supposed emergence of Latino/as in popular music is an issue that provides insight into contemporary issues in politics and relevant societal questions and, as I argue here, provides further insight into questions of cultural identity.

All of this establishes the rationale for this research and the questions that follow: What else can be gleaned from the notable presence of Latino/as in US popular culture? Can Latino/as in popular music aid better understanding of Latino/as, Latino/a identity, or, perhaps, the complications of Latino/a identity discourse(s)? In short, one of the aims of this chapter is to reiterate the fact that popular music should not be dismissed as just popular music but is, in fact, an important cultural site of discourse, debate, and conflict. Thus, a premise of this chapter is that some of the tensions and complications of Latino/a identity are articulated in media and popular culture. At the same time, some further

7 Guzmán and Valdivia, 'Brain, Brow, and Booty: Latina Iconicity in U.S. Popular Culture,' 208–9.
8 The 2000 Census uses the category of 'Spanish/Hispanic/Latino.'
9 For more on Proposition 187, see Hasian and Delgado 'The Trials and Tribulations of Racialized Critical Rhetorical Theory: Understanding the Rhetorical Ambiguities of Proposition 187,' and Ono and Sloop, *Shifting Borders: Rhetoric, Immigration, and California's Proposition 187*.

# Optical Character Recognition (OCR)

Optical Character Recognition (OCR) is a technology that converts images of text into actual digital text.

Once OCR is applied:

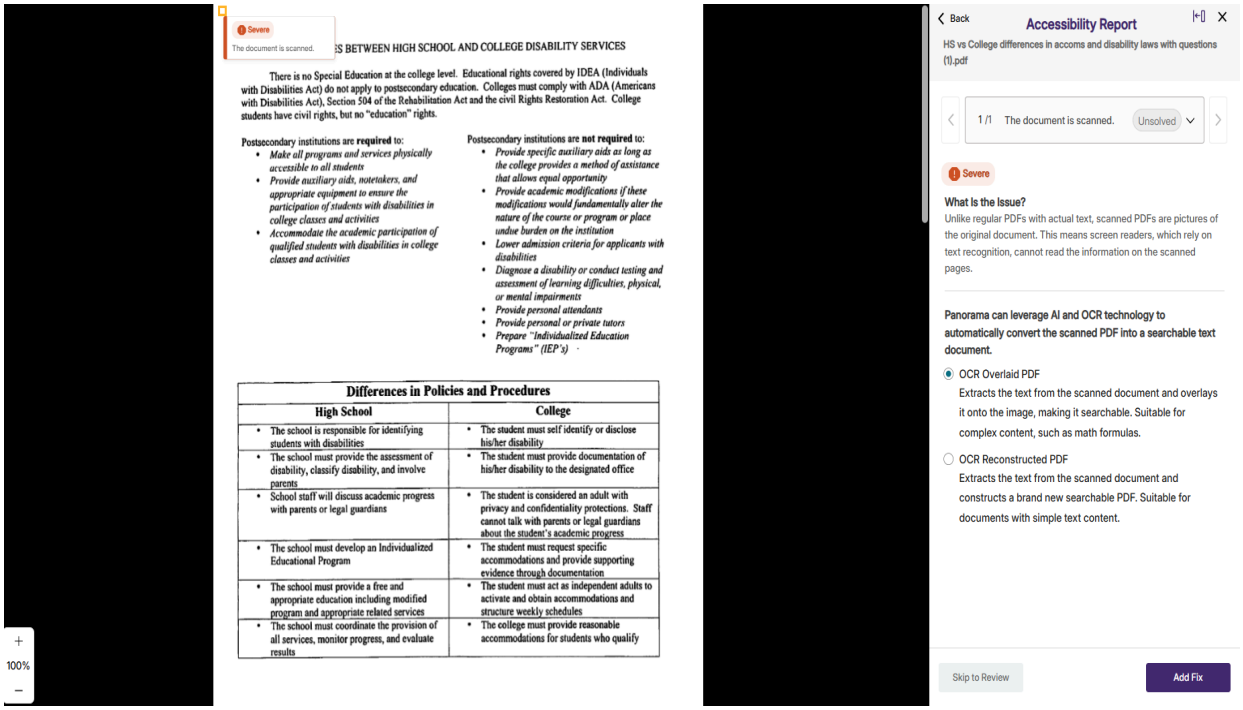
- The text becomes selectable, searchable, and readable by technologies.
- Document elements and structure can be recognized and tagged (headings, paragraphs, tables).

OCR does not automatically create perfect structure, it only turns the visual text into digital text. Structural tagging still needs review and adjustment.

# How YuJa Panorama Can Resolve A Scanned PDF

Follow these steps to apply OCR:

1. **Upload or Locate the Scanned PDF to D2L**
    - Find an existing file or place the PDF in your course content.
2. **Open the PDF in YuJa Panorama**
    - Select the accessibility icon next to the PDF name within D2L content or locate the file using the Panorama course report.
    - Navigate to the Accessibility Report → Structural Remediation Max.
3. **Run Text Recognition (OCR)**
    - If “The document is scanned” appears under Review Issues, then select Fix Issue. If this issue does not appear, the document has likely already been scanned with OCR.
4. **Select Option to Convert**
    - OCR Overlaid PDF - will place an overlay on the image and make it searchable while preserving the original look of the document. This is the recommended option to select.
    - OCR Reconstructed PDF - will pull the text and elements from the document to make a new searchable PDF.
    - Select Add Fix, Proceed to Review, and then Apply Fixes.
    - Depending on the size of the file it may take a moment for the document to process.
5. **Review Tags and Reading Order**
    - Before the document is fully accessible, review the tags and the reading order - this is covered in the next sections of this guide.




---

# Key Points About PDF Scans and Text Recognition (OCR)

- Always check that text in a PDF can be highlighted and copied. If not, then the PDF needs to have OCR applied to it.
- Start with a clean scan for best results:
    - Use a clear original document.
    - Avoid shadows, folds, handwritten marks, or tilted pages.
    - Set the scanner to 300 DPI or higher and black & white or grayscale for text-heavy documents.
    - Save as PDF, not JPEG or PNG.
- If possible, use cleaner alternatives instead of scanning. Many articles and text files are available as accessible links through the library's databases. These can sometimes be recreated in Word or Google Docs and then exported as tagged

PDFs. This typically produces a much higher-quality result than scanning papers or pages of existing documents.

# Structural Elements

# Structural Elements

# Structural Elements

## PDF Tags

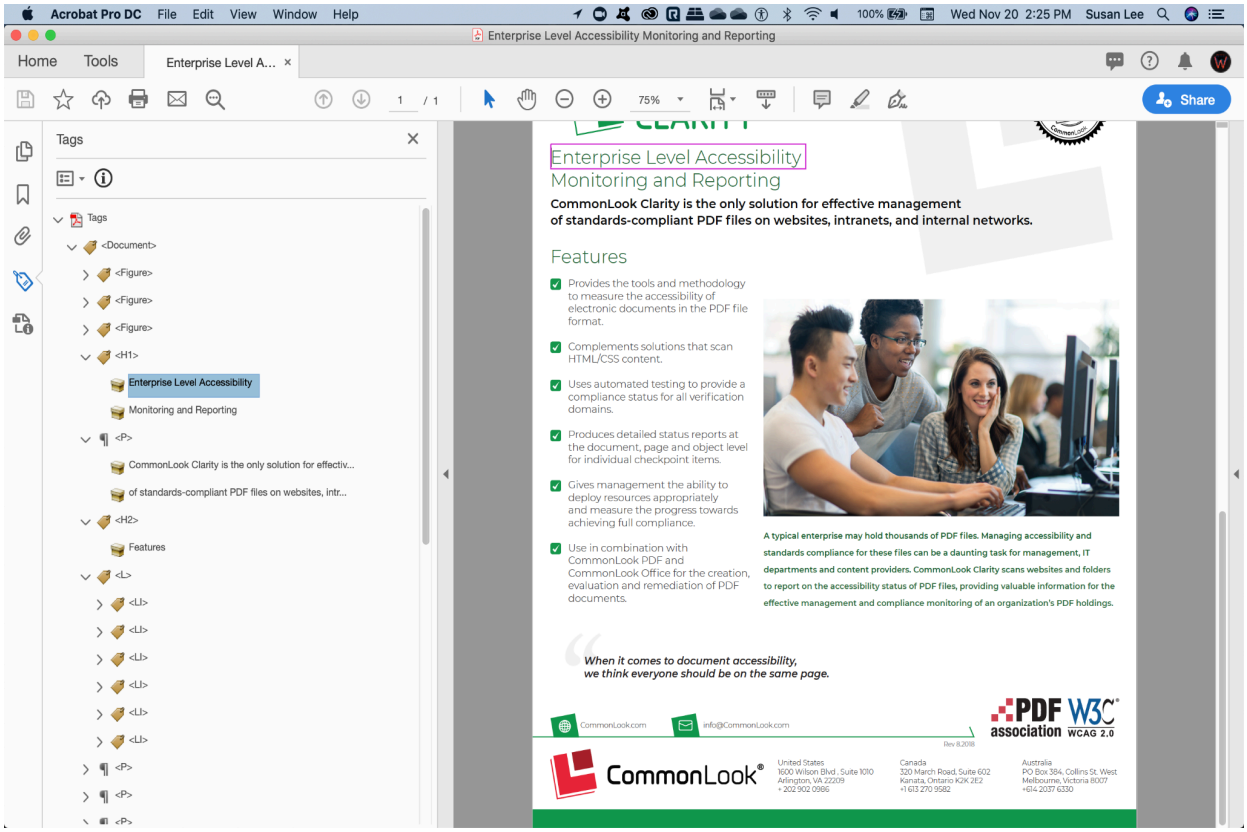
Tags are the structural backbone of a document and help identify various types of content present throughout it. A heading, paragraph, list, table, image, or link are all structural elements that can be tagged so that assistive technologies can interpret and navigate each of them correctly.

These tags also provide assistive technology users the ability to quickly navigate to relevant information. Keyboard commands can be utilized to skip to specific headings, links, or other specialized content.

Think of tags as the “behind-the-scenes outline” that gives meaning to the content. For example:

- Headings help readers navigate quickly between sections.
- Paragraphs mark regular text content.
- Lists and tables organize information logically.
- Figures include alternative text so visuals are described aloud.

When a PDF lacks proper tags, it becomes difficult to navigate with a screen reader, even if all the text is visible to sighted users.




# The Tags Tree

A tag tree is the structural map of a PDF that shows how content is organized and identified for assistive technologies. It is called a “tree” because the tags form a hierarchical, branching structure that usually starts with a top-level <Document> tag and breaks into smaller elements beneath it. Tags are nested to show relationships between elements—for example, a list tag <L> contains list item tags <LI>, and a table <Table> contains rows <TR> with header or data cells (<TH> or <TD>). This nesting indicates both the type of content and the order in which it should be read. A properly structured tag tree ensures elements like headings, paragraphs, tables, and images are accurately tagged and logically organized, which is essential for screen readers to interpret the document correctly.

In Structural Remediation Max, selecting a tag in the tree highlights the corresponding content in the page display. Reviewing the tag tree ensures each page element has a correct tag associated with it—for example, headings are not marked as paragraphs, list items are nested within a list structure, tables use proper header and data cell tags, and images include alternative text when necessary.

# Reviewing PDF Tags

Structural Remediation Max provides the ability for AI to automatically apply accessibility tags within a document.

AI-generated tags still need human review. Check the elements panel after applying tags as some tags may need to be:

- **Changed** - a tagged element was assigned an incorrect element type
- **Removed** - decorative lines or unnecessary elements incorrectly tagged
- **Added** - a page element is not displaying an associated tag

There is an extensive list of element types that Structural Remediation Max provides. While each tag has a function that can increase the accessibility of the document, there are roughly *six core tags* to focus on. Ensuring that these tags are present within the document can help ensure accessibility.

- **Paragraph (P)** – body text
- **Headings (H1–H6)** – “titles” of sections and subsections
- **Lists (L, LI, Lbl, Lbody)** – organized information
- **Tables (TH, TD)** – structured data
- **Links** – clickable resources
- **Figures (with Alt Text)** – images and diagrams

Other tags may be applied, but everything else is for specialized or advanced content. If Structural Remediation Max applies any specialized tags, there is often a core tag nested within them. The 6 core tags are the most essential for an accessible PDF.

---

# How YuJa Panorama Helps with PDF Tagging

Follow These Steps to Apply and Adjust PDF Tags:

1. **Open the PDF in YuJa Panorama**
    - In D2L, select the accessibility icon next to the PDF or access it through the Panorama Course Report.
    - Click Accessibility Report and then Launch Structural Remediation Max.

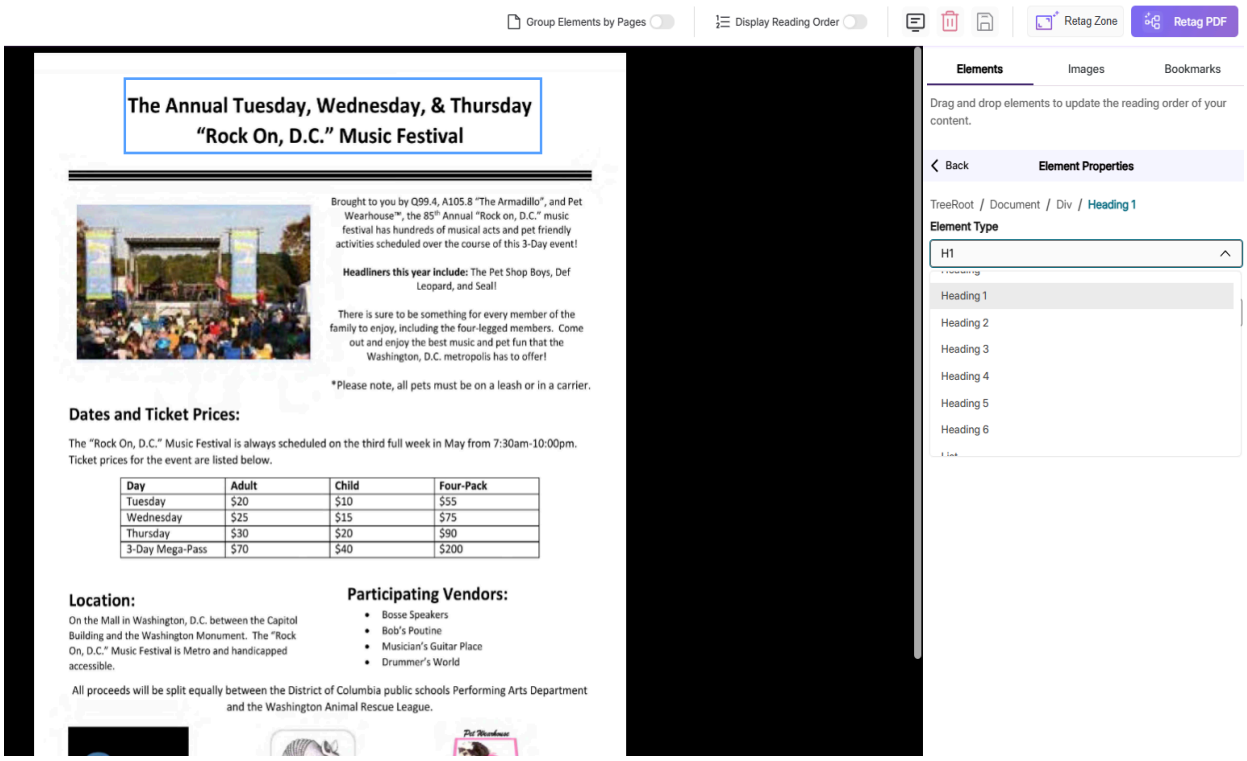
2. **Access the Tags Panel**
    - Once the PDF opens, select the Tags Panel from the right-hand menu. This panel displays the tag tree, which outlines the structure of the document in a nested format.

3. **Review Existing Tags**
    - Open each section of the tag tree by selecting the drop-down arrows. Review the type and nesting of tags to confirm that core structural elements such as headings, paragraphs, lists, tables, and figures are correctly identified.

4. **Identify Missing or Incorrect Tags**
    - Look for untagged elements or tags that do not match the type of content (for example, a heading tagged as a paragraph).
    - Common issues to check include:
        - Improper tags for headings in a logical order <H1>, <H2>, or <H3>
        - Images tagged as text instead of figures - figures/images should have appropriate alternative text
        - Lists without proper structure
            1. <L> - whole list
            2. <LI> - list item
            3. <Lbl> - label such as a bullet point, letter, or number
            4. <Lbody> - text or information within the list item
        - The table is missing header (<TH>) tags, which identify column or row titles. Without them, screen readers cannot convey how the data (<TD>) is organized, making the table difficult to interpret.

5. **Add or Adjust Tags as Needed**
    - Use the Retag Zone or change the tagged Element Type to create or correct tags.
        - To tag untagged content, select Retag Zone or Retag PDF in the upper right corner of Structural Remediation's interface.
        - To adjust hierarchy, click and drag elements in the tag tree to ensure proper nesting (for example, list items inside a list).
        - Remove unnecessary elements by clicking the three dots after the element in the tag tree and selecting Remove Element.

6. **Apply and Save Changes**
    - When all corrections are complete, click Save to apply the changes. The updated PDF will automatically replace the previous version in D2L with improved accessibility.




---

# All Available Tagged Element Types

For reference, each of the available tags in Structural Remediation Max is listed below. These are not essential for every PDF, but are added to provide context for why some elements may be applied when scanned with the tool. Selecting the heading drop down will show the tag types that correspond with that category.

Items with asterisks next to the name indicate core structural elements to look for or include within the document.

## Document Structure Tags

### Part

Used for large divisions within a document, such as "Part I" or "Part II." Common in long reports or books.

### Sect (Section)

Represents a section of content, usually grouped under a heading. Often created automatically when you tag headings.

### Div (Division)

A generic container that groups related content that does not fit another tag type. Example: a box of instructor notes.

### NonStruct (Non-Structural)

Marks decorative or layout-only content such as a line, border, or space.

### Private

Reserved for software-specific tagging structures. You can ignore this tag.

---

## Text Content

### *Paragraph (P)

Used for normal body text grouped as a paragraph.

### BlockQuote

Marks a long, standalone quote often formatted differently than surrounding <P> text.

### Quote

Used for short quotations within a sentence.

### Caption

Text shown in the document that describes a figure, image, or table.
Example: “Figure 1. The structure of a water molecule.”

### Span

Marks a short portion of text that has special meaning or formatting, such as an emphasized phrase.

### Code

Used to tag snippets of computer code so assistive technology reads them correctly.

---

## Headings

Headings create a logical outline of your document. They allow readers to navigate quickly through sections. Utilizing every heading is not always necessary but the majority of documents should include at least a heading level 1.

***Heading 1 (H1): The main title of the document**

Example: Biology 101: Introduction to Life Sciences

**Heading 2 (H2): Major sections within the document**
Example: Cell Structure and Function

**Heading 3 (H3): Subsections under H2 headings**
Example: Mitochondria and Energy Production

**Heading 4–6 (H4–H6): Additional nested levels, if needed**
Example: ATP Synthesis Pathways

Always move in order (H1 → H2 → H3). Don't skip levels.

---

## Lists

Lists help organize related items and ensure screen readers announce the correct structure.

### *List (L)

The overall container for a set of items.
Example:

1. Measure 50 mL of water.
2. Add 5 g of salt.
3. Stir until dissolved.

### *List Item (LI)

Each bullet or numbered item in a list.
Example:

1. Measure 50 mL of water.

### Label (Lbl)

The bullet or number itself (e.g., "1." or "•").
Example:

"1."

### List Body (LBody)

The text or content of each list item.
Example:

Measure 50 mL of water.

---

## Tables

Tables must be tagged so that screen readers know how data is organized.

### *Table

The overall container that holds all table content.

### *Table Row (TR)

Each horizontal line of data.

### *Table Header (TH)

A header cell that labels the entire row or column.

### *Table Data (TD)

A regular data cell.

### Table Head (THead)

Groups the header rows together (optional).

### Table Body (TBody)

Groups the main data rows.

### Table Foot (TFoot)

Groups footer or summary rows (rarely used).

---

## References and Notes

### Note

Marks a footnote or endnote.
Example: “[1]Data collected from NOAA Climate Report (2024).”

### Reference

Used for in-text citations or cross-references.
Example: “(See Table 2 for comparison.)”

### BibEntry

A single entry in a bibliography or reference list.
Example: “Smith, J. (2023). Climate Change Impacts on Agriculture. Springer.”

### *Link

Tags a hyperlink or email address.
Example: Elgin Community College Accessibility Resources

### Annot

An annotation or comment on the PDF, such as an instructor note.
Example: “Please expand this section.”

---

## Visual and Formula Content

### *Figure

Used for images, diagrams, or charts. Always add meaningful alt text.
Example: “Bar chart comparing energy output of solar, wind, and hydro sources.”

### Formula

Used for mathematical or chemical expressions.
Example: “$x_{12} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$” or “2H2O→2H2+O2”

### Art

Represents a complex or grouped graphic (like a diagram made of multiple shapes). Figure tags are more commonly used.

### Artifact

Used to mark decorative or background elements that screen readers should ignore, such as a line, logo, or page number. If items are incorrectly identified in a document such as a fold, empty space, or page speckle these tags can be removed instead of marked as an artifact.

---

## Specialized or Rare Tags

### TOC (Table of Contents)

Marks the entire table of contents section.

### TOCI (Table of Contents Item)

Marks each entry in the table of contents.
Example: “Chapter 1: Introduction - Page 3.”

### Index

Used for an index section.
Example: “Accessibility - 45; PDF Tags - 56.”

### Ruby, RB, RT, RP

Used in East Asian texts for pronunciation guides.
Example: 東京 (とうきょう / Tokyo)

### Warichu, WT, WP

Used for inline commentary in Japanese text.

### Form

Used for interactive form fields.
Example: Name: [ ] Email: [ ]

---

# Key Points About Tagged Structural Elements

- Tagged structural elements make documents easier to navigate and understand using assistive technology.
- Tags that are automatically applied should be reviewed and may need editing. PDFs with complex structure may require time to review and revise.
- If elements are missed in the scan you may need to retag the document or use the retag zone feature.
- Structural Remediation Max has an extensive list of tagged element types that may not necessarily be essential for PDF accessibility. Core tags may be nested within these specialized tags, and include:
    - Paragraph (P) – body text
    - Headings (H1–H6) – “titles” of sections and subsections
    - Lists (L, LI, Lbl, Lbody) – organized information
    - Tables (TH, TD) – structured data
    - Links – clickable resources
    - Figures (with Alt Text) – images and diagrams

If a document contains highly specialized content, or there are difficulties using the tool please feel free to reach out to an Instructional Technology & Accessibility Coordinator for support.


- Ryan Karp 847-214-7101 rkarp@elgin.edu
- Kyle Ahlmann 847-214-7476 kahlmann@elgin.edu

# Reading Order


# Reading Order

## What “Reading Order” Means

Reading order determines the sequence in which a screen reader or assistive technology reads the content presented in a document. It should follow the natural, logical flow of the page in the order that every reader uses.

If the reading order is out of place, a screen reader may jump from a footer to a title, then to a sidebar, and back to the main text, making the document confusing or even unusable.

After tagged structural elements have been applied and reviewed, the reading order should be checked and adjusted as needed.

## Adjusting Reading Order

Structural Remediation Max provides the ability for users to review and edit the assigned reading order of each tagged structural element.

This can be done by hovering the mouse over each element and gradually moving from top to bottom to check each tag in the elements panel or by toggling “Display Reading Order” on from the top menu of Structural Remediation Max. With this enabled, a number will appear in the upper left corner of each tagged element’s bounding box shown in the PDF display. This number will correspond to the reading order shown in the elements panel. If there is anything that is out of order, adjust it manually by dragging and dropping the element to the correct spot in the elements list.

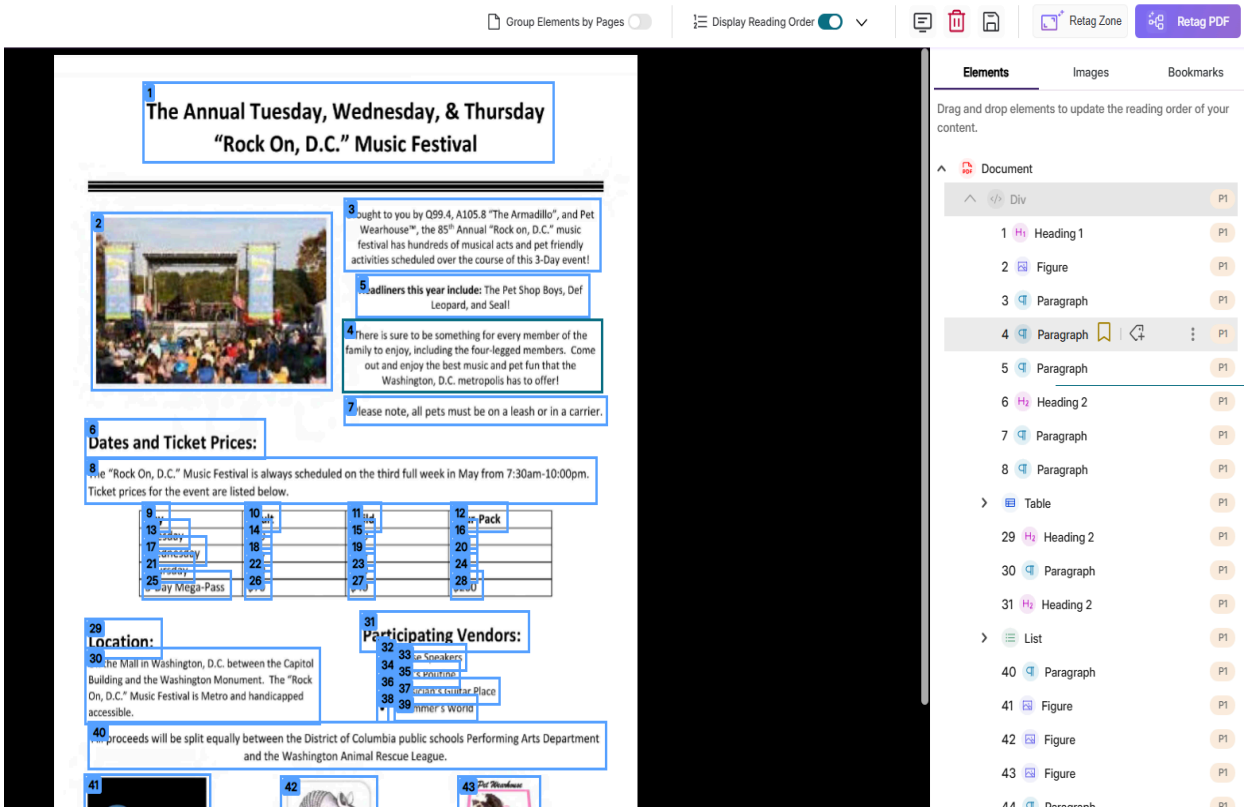

# How YuJa Panorama Helps Fix Reading Order

Follow These Steps to Adjust Reading Order:

1. **Open the PDF in YuJa Panorama**
    - In D2L, select the accessibility icon next to the PDF or access it through the Panorama Course Report.
    - Click Accessibility Report and Launch Structural Remediation Max.

2. **Go to the Elements Panel**
    - Open each tab by selecting the drop down arrow until a number is displayed next to the tab
    - If "Display Reading Order" is toggled off from the upper menu, use the mouse to hover over each element starting at the top and moving down to verify that the sequence of bounding boxes shown in the PDF display is accurate.
    - If "Display Reading Order" is toggled on from the upper menu, a number will appear in the upper left corner of each bounding box, showing the order in which elements will be read. This number can be matched to the number next to the element types listed in the elements panel.

3. **Review the Current Reading Sequence**
    - Check if each item follows the correct flow of content.
    - Watch for common issues such as:
        - Titles being read after the body text
        - Content that is placed side by side ordered incorrectly
        - Information within tables reading in rows instead of columns
4. **Adjust the Reading Order as Needed**
    - Click and drag elements in the reading order list to reorganize them
5. **Apply and Save Changes**
    - Click Save to apply fixes




---

# Key Points About Reading Order

- Reading order should match the visual order of the page's intended flow and sequence.

- Display Reading Order can be toggled on in the upper menu to show all reading order numbers in every bounding box. Using a mouse to hover over each element in the element panel will also show the location and order of each tagged element in the document.
- Multi-column layouts, tables, text boxes, sidebars, or pages with images and captions are most common for presenting reading order issues.
- Order can be adjusted by dragging and dropping the elements in the elements panel.