# Automation in PDF Accessibility: What Can and Can't be Done



#### A COMMONLOOK PRESENTATION



## Agenda

## Automation in PDF Accessibility: What Can and Can't be Done



- The Need For, and the Benefits Of, Accessible PDF
- Regulations and Standards
- Automation in Testing and Reporting
- Automation in PDF Remediation
- Automation in Accessible PDF Creation
- Parting Thoughts (and Tips)
- Comments, Questions, and Answers

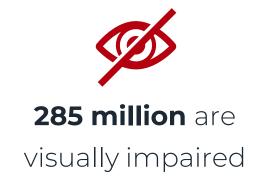


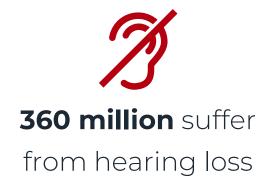
## Over 1 billion<sup>1</sup> people live with some form of disability





**1 in 7** people or 15% of the world's population







Disability is part of the

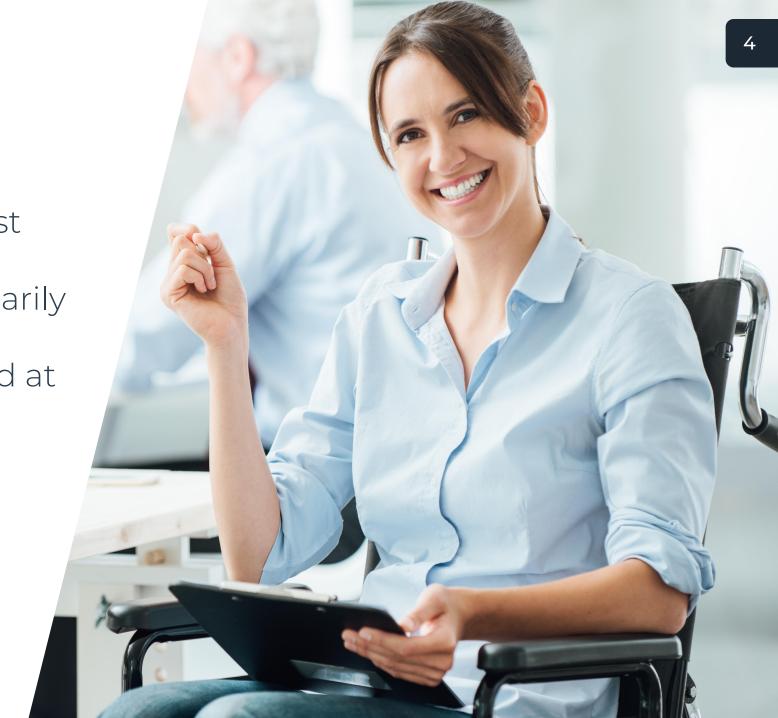
human condition. Almost

everyone will be temporarily

or permanently impaired at

some point in life.





## Accessibility Regulations Demand Compliance



- Section 504 Programs/ Activities receiving Federal Funding
- Section 508 (US Federal Government)
- ADA
  - ▶ Title I Employees
  - ▶ Title II State and Local Government
  - ► Title III "Places of Public Accommodation" includes "Places of Education"
    - Includes websites and electronic documents





## Accessibility Standards for PDF



- Web Content Accessibility Guidelines (WCAG) 2.0
  - WCAG 2.0AA (the "new" Section 508)
  - ► HHS (maps to WCAG 2.0AA)
  - ▶ WCAG 2.1 released in June, 2018.
    - Not much really applies to PDF creation/ remediation
- PDF/UA
  - PDF Association
  - ► ISO 14289-1
  - PDF Specific





## Other Benefits or Uses of Tagged PDF



#### Tagging in PDFs allows for:

- Making them accessible for people with disabilities and using assistive technologies\*
  - Reduce legal risk of litigation due to inaccessible electronic content
- Copying and pasting text and graphics into other applications\*
- Automatic reflow of text and associated graphics to fit a page of a different size than was assumed for the original layout\*
- Processing text for such purposes as searching, indexing, and spell-checking\*
- Conversion to other common file formats (such as HTML, XML, and RTF) with document structure and basic styling information preserved\*

\*From ISO 32000-1-2008 and ISO 32000-2:2017 section 14.8.1.



## Spoiler Alert!



PDF accessibility can NOT be done 100% through automation (regardless of the step you're in)!



## Document Creation Creating Individual PDFs From the Source



#### Authoring in Word, PPT, InDesign, other:

- Use automated checkers built into authoring tools.
  - Caution Most don't test against standards!
- Being proactive saves the time and money otherwise spent on remediation.
- It's generally easier to fix issues during creation as opposed to doing PDF remediation!
- TEST the output against standards/ fix remaining issues!

#### **Document Accessibility Basics**

- ✓ Reading order & formatting
- ✓ Using color, italics, symbols
- ✓ Impact of contrast
- ✓ Headings & styles, alternate text for images, data tables

#### **Accessibility Tools**

- ✓ Microsoft Word & PowerPoint built in accessibility checkers
- ✓ Programs & plugins (e.g., MadeToTag for InDesign, CommonLook Office for Word and PowerPoint)



## Document Creation – Part 2: Creating PDFs from a Template and Data Source



- Another "hot button" topic!
  - What's "trending" in PDF accessibility!?
- Useful for "bulk" or "volume" documents.
- Templates are key:
  - It \*can\* work \*if\* things are set up correctly You have to KNOW what you're doing!
  - ▶ It's automated \*after\* the (manual) template setup phase
  - No template? Not setup correctly? You're going to have problems!
- Outsourcing? Choose the tools/ people who KNOW what they're doing!
- TEST the output against the standard(s)!

#### **Document Accessibility Basics**

- ✓ Reading order & formatting
- ✓ Using color, italics, symbols
- ✓ Impact of contrast
- ✓ Headings & styles, alternate text for images, data tables

#### **Proven Tools for This Approach**

✓ CommonLook Dynamic



## Testing for Accessibility:

What Automation CAN and CAN'T Do...



#### Automation CAN\*:

- Check for Tags (presence)
- Test for headings/ proper "nesting"
- Verify presence of alt text on Figures/ Links
- Verify presence of tooltips on form fields
- Verify all annotations/ form fields are tagged
- Identify/ "flag" color use/ changes
- Test for table regularity, TD and TH are children of TR
- Verify LI, Lbl, and LBody "Parenting"

#### Automation CAN'T\*:

- Verify semantic appropriateness of all tags
- Verify accuracy of artifacted content
- Verify correct reading order
- Verify accuracy of alt text/tooltips
- Identify color/ contrast issues
- Verify accurate metadata
- Verify correct associations in complex tables

#### \* Generally Speaking – Tool Functionality May Vary



## Testing Tools



- Single-Document Testing
  - Adobe Acrobat,
  - PAC 3, and/or
  - CommonLook Validator (or CommonLook PDF).
- In bulk
  - CommonLook Clarity
- Doing it internally? Beware the learning curve!

- Thoughts on Scanning/ Outsourcing:
  - Good for:
    - Discovering then Testing if locations are unknown.
    - Assessing in bulk
    - When time, expertise, knowledge of software/ standards is lacking.
  - Caution: Some HTML checking vendors claim to test PDFs but don't really know PDF



## Remediating PDFs:

### How much remediation can be done with automation?



- Automation Can\*...
  - OCR scanned PDFs
  - Add tags
  - Do a "best guess" on tagging correctly
  - ▶ TRY for accuracy in:
    - Reading order
    - Table, List, TOC tagging

- Automation Can't\*...
  - Always get the tagging right
  - Properly handle complex tables, lists,
     Tables of Contents, reading order...
  - Always provide accurate Alt text for Figures, Links,
  - Always provide accurate Tooltips for form fields
  - ► Fix color or contrast issues



<sup>\*</sup> Generally Speaking – Tool Functionality May Vary

## How much remediation TO do!? (Is "Close" Good Enough?)



#### The Arguments for Automation:

- Using Automation to get "most of the way there" saves time, money, training efforts.
- Getting more docs close faster is better than taking a lot of time to get fewer documents to 100%
- "Close" is better than nothing.

#### Think about the end user:

- If it's not 100% it falls short of the mark and isn't really accessible.
- Getting 60, 70, or even 90% of the way up the ramp doesn't get you in the building.
- "You're not worth it."



## Tools for Manual Remediation: Adobe Acrobat



- Included in Acrobat Pro
- Auto-tag PDFs (single or batch)
  - Bulk process may include additional steps/ accuracy
- Has some automated "Fix" options for some failures/ issues:

- Bulk remove empty tags
- Show Role Mapped tags
- Assign IDs to TH cells
- Automatically generate bookmarks from headings



## Tools for Manual Remediation: CommonLook PDF



- Tests against multiple standards:
   WCAG 2.0, (508), PDF/UA, and HHS
- Check and Fix document for corruption
- Bulk Processes to:
  - Find and Replace symbols, etc.
  - Search Tags/ perform operations
  - Convert Role Mapped tags to Standard PDF tag
  - Cleanup empty tags
  - Create header/ footer pagination artifacts
  - Generate bookmarks based on headings

- Automatically generate Lists, TOC, Tables from selection
- Automatically assign IDs to TH cells
- 1-Click to automatically link data cells to header cells in complex tables
- Tag untagged Annotations in the correct reading order and tag type
- Automatically assign Alt text to links/ Contents to Annotations
- Automatically create Tooltips for Form Annotations
- Fix Wizard automatically corrects most failures and warnings



### CommonLook AI Cloud



- Patent-pending technology.
- Built on:
  - CommonLook's 20 years of experience in PDF accessibility and
  - Coupled with advanced machine learning
- Software learns from existing documents gets even better over time!
- Incredibly fast; Highly accurate; Economical
  - Faster \*and\* more cost effective as compared to traditional manual remediation
- Great for large quantities of documents
- Backed by the CommonLook Guarantee



## Summing It Up:

Considering an Automated Solution?



- Automation in PDF accessibility may have its place.
  - Human interaction is required at some point.
- TEST (including manual checks) to know if the automated work was done correctly!
- Life lessons apply: Anything worth doing, and doing well, is probably going to take some hard work at some point.
- Anything less than 100% isn't accessible, may not be usable, could be insulting, and could create legal risk!





## Thanks for Listening!



## **Paul Rayius**

Director of Training <a href="mailto:prayius@commonlook.com">prayius@commonlook.com</a>

Office: (202) 469-4487

Email info@commonlook.com

Website: commonlook.com

