Mapping based on forensic-workbench

SLIDE #1: Overall summary and overview

- total number of files
- number of files per file type (graph)
- number of files per threat level (ok, medium, high)
- number of sanitized files
- number of quarantined files

REST OF THE SLIDES

- type of data per file type (internal/external hyperlink, metadata, javascript, embedded files, macros) and tags for each of them (allow, disallow, sanitized)
- these results are present per file upload, not as general data – crawler for extracting them

Mapping based on file-drop

Similar output/data as forensic-workbench.

SECTIONS

- FILE ATTRIBUTES: type, size, name
- ACTIVE CONTENT THAT HAS BEEN SANITISED (REMOVED) – metadata, javascript, embedded files... More on - https://glasswallsolutions.com/technology/
- OBJECTS & STRUCTURES THAT HAVE BEEN REPAIRED
- OBJECTS & STRUCTURES THAT ARE UNABLE TO BE REPAIRED

Try to find more on: https://glasswallsolutions.com/wp-content/uploads/2020/01/Glasswall-d-FIRST-Technology.pdf

What gets sanitized (content management policy across supported file types):

pdf	**word**	**ppt**	**xls**
metadata	metadata	metadata	metadata
javascript	macros	macros	macros
acroform	embedded_files	embedded_files	embedded_files
actions_all	review_comments	review_comments	review_comments
embedded_files	external_hyperlinks	external_hyperlinks	external_hyperlinks
external_hyperlinks	internal_hyperlinks	internal_hyperlinks	internal_hyperlinks
internal_hyperlinks	dynamic_data_exchange	embedded_images	dynamic_data_exchange
embedded_images	embedded_images		embedded_images
value_ouside_reasonable_limits			

For tiff extension, there is no policy for sanitization, just that geotiff is allowed.

Each file has specific # of content groups
Each content group has 4 sections:

- Brief description
- Sanitization items
- Remedy items
- Issue items

Sanitization, remedy and issue items are part of the report on file-drop and forensic-workbench.

Technical description is extracted if any of these items has something detected.

If all of these items, across all content groups, have 0 count value – file is clean.

xPath:

- count(//gw:ContentGroup)
- count(//gw:SanitisationItems[@itemCount>0]) – determine if the file was sanitized
- count(//gw:RemedyItems [@itemCount>0]) – determine if the file was remediated/repaired
- count(//gw:IssueItems [@itemCount>0]) – determine if there was smth that could not be repaired
- //gw:FileType/text() – returns type of the file that was processed