

Document Versioning & Control

By Peter Hemphill

My Perspective & Background

- Senior System/Software Designer
- 17 years designing systems to manage documents and related content
- Currently manage systems maintaining over 3 million documents from hardcopy to electronic formats.
- Manage, control and version republished documents
 - Need to take in virtually any kind of document/attachment
 - Need to manage and track changes to those documents
 - Need to handle varying structures, formats and renderings of documents.
 - Need to manage documents electronically (including hardcopy)

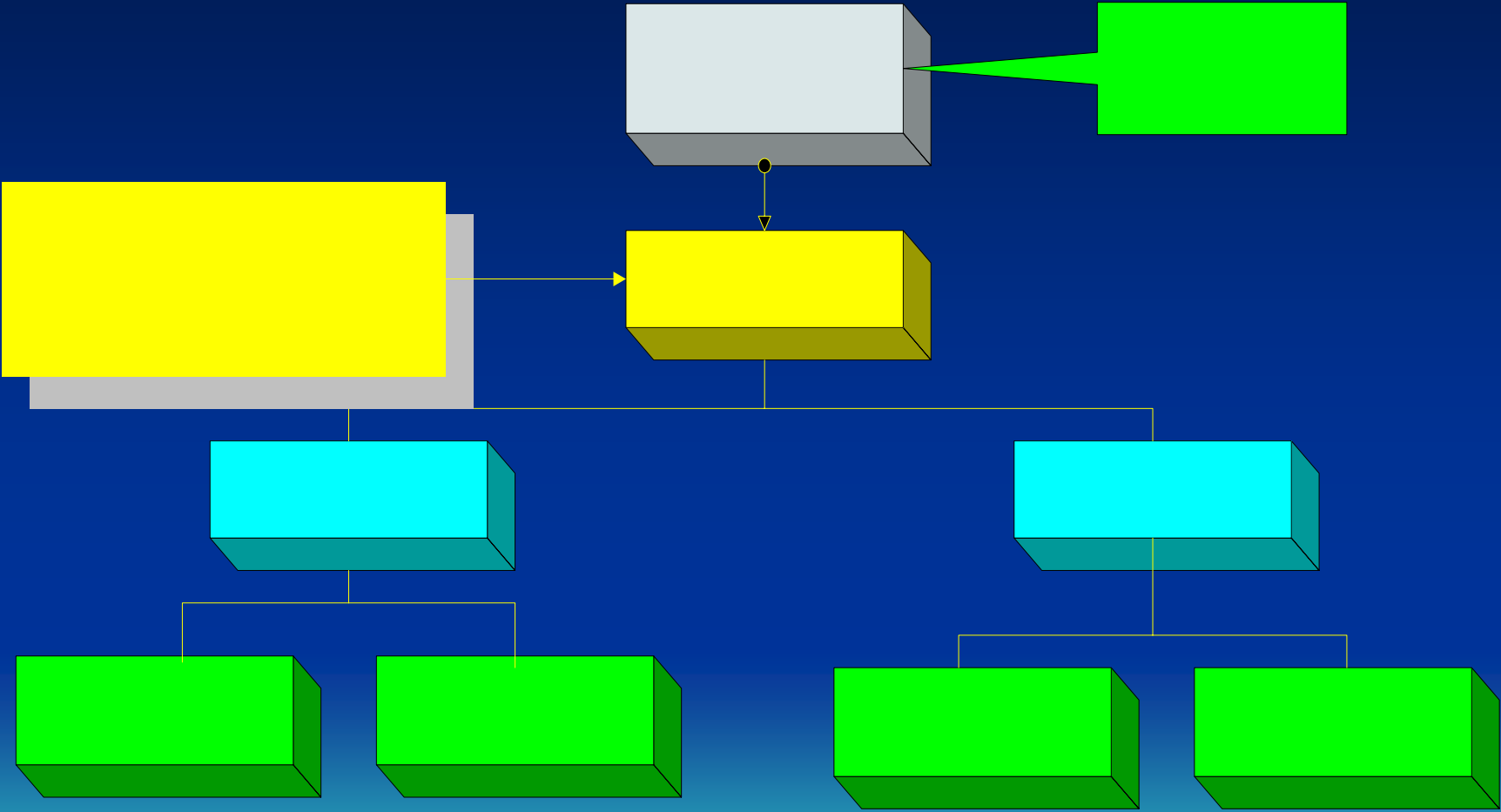
What is “Version Control”

- Version Control is really a combination of:
 - Managing relationships between levels in a hierarchical structure that comprises a document or set of documents.
 - Storing/Managing and using information about the changes (“metadata”).
 - Cross-Referencing to other documents or sets of documents indicating replacements or other applicable documents.

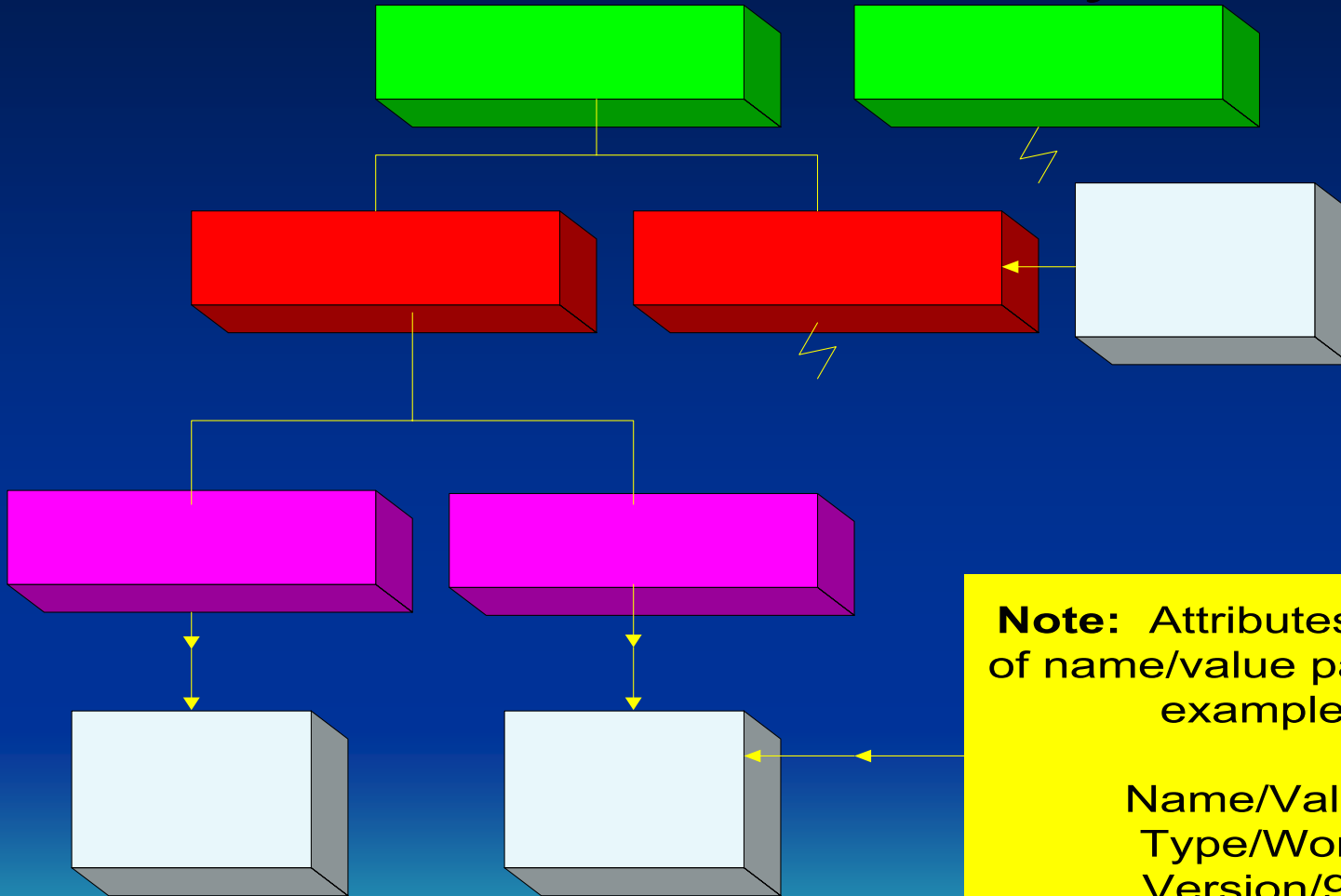
Document Structure

- Documents have a hierarchical “Bill-of-Materials” type structure.
 - Building a car is conceptually like the hierarchy of building documents!
 - Very flexible structure to manage versions from the largest piece down to the page if you like.
 - Each level has related metadata/attributes
 - Who, what, where (location), when and why.
 - Managing “Objects” and attachments as well as hardcopy.

Document Hierarchy Levels



Document Hierarchy Levels



Note: Attributes consist of name/value pairs. For example:

Name/Value
Type/Word
Version/97

Storing, Identifying & Grouping Document Information

- Each level has a unique identifier for each instance
 - Like an ISBN, only for each level/piece of a document that needs to be versioned.
 - Use the level that suits your operation to store in groups that make sense.
 - Attaching an OCR label or page to hardcopy/folder
- Three (or more) methods to file items
 - Just put the items on a shelf one after the other
 - File/Group by a level that makes sense for your needs
 - By document number, subject or other means. This can lead to “interpretation” issues.
- Each level has a relationship to higher and lower levels.

Document Cross-References

- Examples of Cross-References:
 - Superseding
 - Superseded by
 - Use
 - “Refer to”
 - Cancelled by
 - In-Lieu-Of
 - Others
- Cross-References link to different levels in the hierarchy
 - “Super Family”, Family, Revision, Document Levels

Fundamental Questions to Ask

- What can be *thrown away*?
 - Information multiplies
 - Exponential costs with ripple-effects
 - Reduces scale and increases performance
- What are the costs, regulations and justifications for keeping and controlling data elements/documents?
- What information changes?
- What information needs version control?
- What bibliographic/metadata elements are necessary to find and manage the documents?
- Will an “out-of-the-box” solution *really* work?
 - Flexibility & Customization
 - Is the cost/benefit worth customizing the software? Or is it more cost effective to change processes to adapt to the software?

Conclusion

- Thanks for your time!
- I hope this helps in dealing with the challenges we all face in document management!
- Email: Pete.Hemphill@ihs.com