Meeting Minutes
PWG MFD Model Working Group Teleconference

Meeting was called in order at 11am EDT, June 7, 2007, by Peter Zehler.

Attendees:

Hiromasa Akamatsu	Kyocera Mita
Nancy Chen	Oki Data
Lee Farrell	Canon
Walter Filbrich	Samsung
Harry Lewis	Ricoh-IBM
Bill Wagner	TIC
Craig Whittle	Sharp
Peter Zehler	Xerox

Last Teleconference Meeting Minutes Review –
- The meeting minutes for the last teleconference was accepted without change.

Meeting Agenda was Agreed as Follows –
- Review the updated Scan Job Ticket Life Cycle Diagram (page 21)
- Use Cases Discussions

Discussion on Scan Job Ticket Life Cycle Diagram –
- Peter led the group through the discussion of the diagram:
    - o The model diagram now includes some elements from the PWG Semantic Model version 2 such as MFD imaging services and subunits.
    - o Issue #1: The MFD subunits include a scanner. But it is unclear which physical part of a scanner device we are talking about here.
        - ▪ A subunit is not a device. A scanner is a device. In PWG Semantic Model v2 the subunit "scanner" actually means the hardware component under the platen that does scanning of paper document into a digital image.
        - ▪ There should be common UI as a subunit of MFD that is used by all MFD services. Thus the Walk-up User Interface should not be part of Scanner subunit only. But both Scan Service and Scan Walk-up User Interface use MFD subunits. Also Scanner should not be the subject of discussion in this diagram.
        - ▪ Decision: Delete the scanner subunits. The Scan Service "uses" MFD subunits (e.g. MFD UI), and "Scan Walk-up User Interface" is a software module as part of "Scan Service" which also uses MFD subunits.
    - o Issue #2: Should one Scan Job only contains 0..1 Scan Document? Should it contain 0 or more documents?
        - ▪ If a job contains more than one document, there will be requirements for complicated document management such as splitting job into multiple documents.

- Most application today only manage one document per job, the added complexity is hard to justify.
- o Issue #3: Should the scan operation scan an entire document automatically at once or a document page at a time for allowing approval of the scanned page and then assemble multi-page into a single document?
  - In general we should assume scanning all pages into a single document. For scanning from ADF, multiple pages are automatically scanned until finish. For scanning from the platen, scan is done page-by-page, there should be a signal to tell the end of document.
  - We should consider page level granularity in order to signal a page jam at certain page number, similar to print.
  - If we define document information for each page, we could tie each page to a single document just like a print job. You might have text pages color or mono and photo pages scanned, each with different parameters specified, they need to be combined as a single document. This is typical in a billing job. A higher level job parameter construct that consists of page-by-page parameters could be considered in this case. However it's simpler to have one job with parameters specified before scanning multiple pages. Sometimes you could have an instructional sheet for each multi-page document in a multi-document scanning. This could be viewed as a batch of separate jobs and it is always possible to have a transformation process that can build a document from multiple jobs. We should not consider the complexity of allowing different scan parameters at page level.
  - Decision: Use a simple scan model. The Scan Service has single document per job, i.e. the scan job has the same set of scan parameters in Scan Job Ticket for all images in a scanned document. We will allow vendors to extend the scan job ticket to include page-level instructions. In our model, you need to specify a different set of parameters for a different definition of scan task. It is application-specific to specify how scan jobs are to be combined into documents, each job might have different document format. A Scan Document object will be defined that must have page count, document format, …, etc., the information that can be passed down to other jobs, a print job for example.
- o Open Issue : How should metadata be included in the documents which are to be stored in document repository?
  - A repository is a file system which might not have the capability to carry any metadata at document transfer protocol level. But metadata can be added into a Scan Document that's to be sent and stored in repository, or can be in the Scan document object that's maintained by Scan Service. Should we define a Document Repository API to handle metadata or maintain metadata in Scan Service and metadata become part of document format? For the latter we need to define a document format, but currently we already excluded "defining new document format" from the scope of Scan Service. That means we are not going

to define standard document format containing metadata that is interoperable across all Industry scan services. Alternatively the metadata could be stored in a separate file, but then we would have file naming convention issue.

- Decision: We will delay the decision, pending on decision of whether to include "document format" in Scan Service. Right now the Scan Service can output any kind of scan document it is capable of, including PDF document for scan-to-pdf. An implementation is free to populate scan job or application specific information to be included in a Scan Document along with the scanned image data. We will model Scan Document object which consists of metadata and document image data managed by Scan Service that allows metadata to be added to Scan Document.

- **Next Step:**
  - Action Item: Nancy Chen to update the Scan Job Ticket Life Cycle Diagram.
  - Plan face-to-face meeting subject and documents.

  (Note: Pete has reported that he has not received from Mike any Microsoft's intent on WSD, but Xerox has given green light for his contribution to PWG.)

- **Next Teleconference:**
  EDT June 21, 2007, Thursday, 11am-12pm.