

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:53 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

I figured he was one of the 3 humans who think that typing in structure makes sense. Well I guess Seurat thought painting in halftone was painless too.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 1:50 PM  
**To:** Steven Sinofsky  
**Subject:** RE: XML plans

I try! The reason that he mailed you was that I tried to explain to him that Word is a lot easier than SGML editors *intrinsically* because its storage is isomorphic to its layout. But he actually wrote his damn book (a really good XML book) using an SGML based editor and knows SGML inside and out and is resisting this.

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:45 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

Seems like we should work to set the expectations of people like this and how we see these evolving. If all the SGML people think that the world will finally get it, and the average person will be able to tag a document before formatting it they will only be disappointed in XML and how it manifests itself.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 6:38 AM  
**To:** Steven Sinofsky  
**Cc:** Ralf Harteneck  
**Subject:** RE: XML plans

Hey, I agree with you here. Remember. The abstraction of XML/XSL has to show right through to the UI and that in turn makes the product much harder. I actually like the current HTML annotated with XML model contingent on some perf goals and some cleanliness that marco and I agree about. I'd argue for Excel only because essentially they format from an internal store anyway, but for Word you and I are in synch. For Powerpoint, I'd be prepared to listen to a counter-argument, but have to admit I see that one as a data model. Notice that in both Powerpoint and Excel, I'm still not arguing that XSL solves anything, just that since the layout is hardwired into the code, the data can be pure (and even then Steven, your arguments for Excel and Powerpoint being in HTML still seem pretty good if the perf is there).

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Sunday, January 11, 1998 8:49 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

What I'm confused about is why does everyone think that style sheets will work this time around? I don't think XSL+XML fixed any of the inherent problems of SGML+DTDs. People still don't think in terms of styles and structure when they write, and direct formatting always wins. We can automate some things like tagging bullet lists with styles, but even then customers will direct format those after we do the automatic thing. And the automatic stuff might have gone far enough for a while based on feedback and PSS calls from 97.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Sunday, January 11, 1998 3:10 PM  
**To:** Ralf Harteneck  
**Cc:** Steven Sinofsky  
**Subject:** FW: XML plans

I'm working with a writer on XML support. I'm forwarding on to you some comments he made to me which you are welcome to pursue or not as you see fit. I had to explain to him that Office supporting XML doesn't mean that XML is the native format for all text (which is what SGML folks automatically assume). Obviously, you wouldn't be changing your plans at this point even if you drank all of his kool-aid, but if you are interested, here

is a contact for one of your folks to argue the pro's and con's of structured storage for WP. For what it is worth, while I'd make the argument for Powerpoint and Excel, I wouldn't make it personally for Word myself.

All the best

Adam

-----Original Message-----

From: Richard Light [SMTP:richard@light.demon.co.uk]  
Sent: Friday, January 09, 1998 2:40 AM  
To: Adam Bosworth  
Subject: XML plans

Adam,

I enjoyed meeting with you earlier this week, and I'll be in touch soon about the book idea.

In the meantime, I have been thinking over the little bit that you had time to tell me about your plans for XML support, and I must say that I am left with a feeling of concern. This relates primarily to the word processing side of things, but also applies to the storage of spreadsheets.

To put it bluntly, I think that you are in danger of passing up a wonderful opportunity if you don't store Word documents as well-formed XML.

The documents would have to be accompanied by XSL style sheets (one or more per document, each based on an underlying style for the template used).

As you rightly point out, most Word users couldn't care less about the structure of their documents or even the use of styles. However:

- the use of templates for letters, memos, etc. with suitable markup for header information could make them into searchable information resources at no cost to the user whatsoever;
- storing the document as valid XML means that it can be picked up and used by third-party software without any need for conversion or filtering;
- separating out markup from styles makes it much easier to support multiple views of the document within Word (online; paginated), and opens up the possibility of allowing users to fine-tune the styling of each view independently.

The second point is crucial to processes like publication, where a document will typically be authored in Word, and then converted at some point to a DTP package's internal format. Once that happens, any changes to the text will be made in the DTP environment, and not reflected in the 'source' Word document. The value of the out-of-date Word document is thus radically reduced, because it doesn't reflect the printed document correctly. In general terms, this limits the potential re-use of the document (even when all that is required is 'web page and DTP'ed hard copy').

If the Word document is actually XML, and the DTP package can work directly with XML documents and also save them as XML, then any changes made at the DTP stage will still be available from Word.

Storing Word documents as XML also opens up possibilities for search engines and databases to exploit the markup for more precise retrieval, as outlined in my first point.

In a way, you could see XML + XSL as the successor to RTF.

MS/CR 0003120  
CONFIDENTIAL

It seems to me that the biggest challenge to this vision that you face is getting documents well-formed and keeping them that way. (In our meeting we discussed the issue of horrible non-hierarchical HTML pages!)

Incidentally, I also have reservations about the use of HTML as any sort of storage format. It is fine as a delivery medium, and I like the way that XSL lets you deliver HTML - but that is its proper place in the overall scheme of things, in my view.

I could go into more detail (as you may gather, I have strong feelings on the subject!), but I thought it best to give you an initial set of comments and see what you think. Apart from anything else, I may have misunderstood the comments to which this email is a reaction!

I look forward to hearing from you, and getting your advice on how we might take this dialogue forward.

Best wishes,

Richard.

Richard Light  
SGML/XML and Museum Information Consultancy  
richard@light.demon.co.uk

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:56 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

We were not specific at all. We just said we were using XML to maintain the edit state and meta information, and that HTML was used for everything in the traditional way so that browsers could universally display Office documents. I would sometimes say that XML will just ignored by today's browsers and is only used by Office, but in the future browsers, in addition to third parties and other s/w, might choose to do something with the XML.

Basically, we just want people to know we're doing HTML at this point. We threw in the XML because it was magic. Everything is just "directional" for us now, with no specifics on implementation, timeframe, etc.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 1:54 PM  
**To:** Steven Sinofsky  
**Subject:** RE: XML plans

One thing you should tell me how to handle, questions on what it means for Office to be "supporting" XML. How explicit should I be that this doesn't mean the text is all stored in XML, just the semantic annotations?

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:53 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

I figured he was one of the 3 humans who think that typing in structure makes sense. Well I guess Seurat thought painting in halftone was painless too.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 1:50 PM  
**To:** Steven Sinofsky  
**Subject:** RE: XML plans

I try! The reason that he mailed you was that I tried to explain to him that Word is a lot easier than SGML editors *intrinsically* because its storage is isomorphic to its layout. But he actually wrote his damn book (a really good XML book) using an SGML based editor and knows SGML inside and out and is resisting this.

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:45 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

Seems like we should work to set the expectations of people like this and how we see these evolving. If all the SGML people think that the world will finally get it, and the average person will be able to tag a document before formatting it they will only be disappointed in XML and how it manifests itself.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 6:38 AM  
**To:** Steven Sinofsky  
**Cc:** Ralf Harteneck  
**Subject:** RE: XML plans

Hey, I agree with you here. Remember. The abstraction of XML/XSL has to show right through to the UI and that in turn makes the product much harder. I actually like the current HTML annotated with XML model contingent on some perf goals and some cleanliness that marco and I agree about. I'd argue for Excel only because essentially they format from an internal store anyway, but for Word you and I are in synch. For Powerpoint, I'd be prepared to listen to a counter-argument, but have to admit I see that one as a data model. Notice that in both Powerpoint and Excel, I'm still not arguing that XSL solves anything, just that since the layout is hardwired into the code, the data can be pure (and even then Steven, your arguments for Excel and Powerpoint being in HTML still seem pretty good if the perf is there).

MS/CR 0003122  
CONFIDENTIAL

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Sunday, January 11, 1998 8:49 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

What I'm confused about is why does everyone think that style sheets will work this time around? I don't think XSL+XML fixed any of the inherent problems of SGML+DTDs. People still don't think in terms of styles and structure when they write, and direct formatting always wins. We can automate some things like tagging bullet lists with styles, but even then customers will direct format those after we do the automatic thing. And the automatic stuff might have gone far enough for a while based on feedback and PSS calls from 97.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Sunday, January 11, 1998 3:10 PM  
**To:** Ralf Harteneck  
**Cc:** Steven Sinofsky  
**Subject:** FW: XML plans

I'm working with a writer on XML support. I'm forwarding on to you some comments he made to me which you are welcome to pursue or not as you see fit. I had to explain to him that Office supporting XML doesn't mean that XML is the native format for all text (which is what SGML folks automatically assume). Obviously, you wouldn't be changing your plans at this point even if you drank all of his kool-aid, but if you are interested, here is a contact for one of your folks to argue the pro's and con's of structured storage for WP. For what it is worth, while I'd make the argument for Powerpoint and Excel, I wouldn't make it personally for Word myself.

All the best  
Adam

-----Original Message-----

**From:** Richard Light [SMTP:richard@light.demon.co.uk]  
**Sent:** Friday, January 09, 1998 2:40 AM  
**To:** Adam Bosworth  
**Subject:** XML plans

Adam,

I enjoyed meeting with you earlier this week, and I'll be in touch soon about the book idea.

In the meantime, I have been thinking over the little bit that you had time to tell me about your plans for XML support, and I must say that I am left with a feeling of concern. This relates primarily to the word processing side of things, but also applies to the storage of spreadsheets.

To put it bluntly, I think that you are in danger of passing up a wonderful opportunity if you don't store Word documents as well-formed XML.

The documents would have to be accompanied by XSL style sheets (one or more per document, each based on an underlying style for the template used).

As you rightly point out, most Word users couldn't care less about the structure of their documents or even the use of styles. However:

- the use of templates for letters, memos, etc. with suitable markup for header information could make them into searchable information resources at no cost to the user whatsoever;

- storing the document as valid XML means that it can be picked up and used by third-party software without any need for conversion or filtering;

**MS/CR 0003123  
CONFIDENTIAL**

- separating out markup from styles makes it much easier to support multiple views of the document within Word (online; paginated), and opens up the possibility of allowing users to fine-tune the styling of each view independently.

The second point is crucial to processes like publication, where a document will typically be authored in Word, and then converted at some point to a DTP package's internal format. Once that happens, any changes to the text will be made in the DTP environment, and not reflected in the 'source' Word document. The value of the out-of-date Word document is thus radically reduced, because it doesn't reflect the printed document correctly. In general terms, this limits the potential re-use of the document (even when all that is required is 'web page and DTP'ed hard copy').

If the Word document is actually XML, and the DTP package can work directly with XML documents and also save them as XML, then any changes made at the DTP stage will still be available from Word.

Storing Word documents as XML also opens up possibilities for search engines and databases to exploit the markup for more precise retrieval, as outlined in my first point.

In a way, you could see XML + XSL as the successor to RTF.

It seems to me that the biggest challenge to this vision that you face is getting documents well-formed and keeping them that way. (In our meeting we discussed the issue of horrible non-hierarchical HTML pages!)

Incidentally, I also have reservations about the use of HTML as any sort of storage format. It is fine as a delivery medium, and I like the way that XSL lets you deliver HTML - but that is its proper place in the overall scheme of things, in my view.

I could go into more detail (as you may gather, I have strong feelings on the subject!), but I thought it best to give you an initial set of comments and see what you think. Apart from anything else, I may have misunderstood the comments to which this email is a reaction!

I look forward to hearing from you, and getting your advice on how we might take this dialogue forward.

Best wishes,

Richard.

Richard Light  
SGML/XML and Museum Information Consultancy  
richard@light.demon.co.uk

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 2:02 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

That would be fine. We won't be adding anything to this message. The next directional thing will be on collaboration, which will be "enabled" by HTML/XML. We'll talk about this at DEMO/Internet Showcase.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 1:58 PM  
**To:** Steven Sinofsky  
**Subject:** RE: XML plans

And I assume that this is what you want me to say as well.

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:56 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

We were not specific at all. We just said we were using XML to maintain the edit state and meta information, and that HTML was used for everything in the traditional way so that browsers could universally display Office documents. I would sometimes say that XML will just be ignored by today's browsers and is only used by Office, but in the future browsers, in addition to third parties and other s/w, might choose to do something with the XML.

Basically, we just want people to know we're doing HTML at this point. We threw in the XML because it was magic. Everything is just "directional" for us now, with no specifics on implementation, timeframe, etc.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 1:54 PM  
**To:** Steven Sinofsky  
**Subject:** RE: XML plans

One thing you should tell me how to handle, questions on what it means for Office to be "supporting" XML. How explicit should I be that this doesn't mean the text is all stored in XML, just the semantic annotations?

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:53 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

I figured he was one of the 3 humans who think that typing in structure makes sense. Well I guess Seurat thought painting in halftone was painless too.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 1:50 PM  
**To:** Steven Sinofsky  
**Subject:** RE: XML plans

I try! The reason that he mailed you was that I tried to explain to him that Word is a lot easier than SGML editors *intrinsically* because its storage is isomorphic to its layout. But he actually wrote his damn book (a really good XML book) using an SGML based editor and knows SGML inside and out and is resisting this.

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Monday, January 12, 1998 1:45 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

Seems like we should work to set the expectations of people like this and how we see these evolving. If all the SGML people think that the world will finally get it, and the average person will be able to tag a document before formatting it they will only be disappointed in XML and how it manifests itself.

**MS/CR 0003125**  
**CONFIDENTIAL**

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Monday, January 12, 1998 6:38 AM  
**To:** Steven Sinofsky  
**Cc:** Ralf Harteneck  
**Subject:** RE: XML plans

Hey, I agree with you here. Remember. The abstraction of XML/XSL has to show right through to the UI and that in turn makes the product much harder. I actually like the current HTML annotated with XML model contingent on some perf goals and some cleanliness that marco and I agree about. I'd argue for Excel only because essentially they format from an internal store anyway, but for Word you and I are in synch. For Powerpoint, I'd be prepared to listen to a counter-argument, but have to admit I see that one as a data model. Notice that in both Powerpoint and Excel, I'm still not arguing that XSL solves anything, just that since the layout is hardwired into the code, the data can be pure (and even then Steven, your arguments for Excel and Powerpoint being in HTML still seem pretty good if the perf is there).

-----Original Message-----

**From:** Steven Sinofsky  
**Sent:** Sunday, January 11, 1998 8:49 PM  
**To:** Adam Bosworth  
**Subject:** RE: XML plans

What I'm confused about is why does everyone think that style sheets will work this time around? I don't think XSL+XML fixed any of the inherent problems of SGML+DTDs. People still don't think in terms of styles and structure when they write, and direct formatting always wins. We can automate some things like tagging bullet lists with styles, but even then customers will direct format those after we do the automatic thing. And the automatic stuff might have gone far enough for a while based on feedback and PSS calls from 97.

-----Original Message-----

**From:** Adam Bosworth  
**Sent:** Sunday, January 11, 1998 3:10 PM  
**To:** Ralf Harteneck  
**Cc:** Steven Sinofsky  
**Subject:** FW: XML plans

I'm working with a writer on XML support. I'm forwarding on to you some comments he made to me which you are welcome to pursue or not as you see fit. I had to explain to him that Office supporting XML doesn't mean that XML is the native format for all text (which is what SGML folks automatically assume). Obviously, you wouldn't be changing your plans at this point even if you drank all of his kool-aid, but if you are interested, here is a contact for one of your folks to argue the pro's and con's of structured storage for WP. For what it is worth, while I'd make the argument for Powerpoint and Excel, I wouldn't make it personally for Word myself.

All the best

Adam

-----Original Message-----

**From:** Richard Light [SMTP:richard@light.demon.co.uk]  
**Sent:** Friday, January 09, 1998 2:40 AM  
**To:** Adam Bosworth  
**Subject:** XML plans

Adam,

I enjoyed meeting with you earlier this week, and I'll be in touch soon about the book idea.

In the meantime, I have been thinking over the little bit that you had time to tell me about your plans for XML support, and I must say that I am left with a feeling of concern. This relates primarily to the word processing side of things, but also applies to the storage of spreadsheets.

To put it bluntly, I think that you are in danger of passing up a

**MS/CR 0003126**  
**CONFIDENTIAL**



wonderful opportunity if you don't store Word documents as well-formed XML.

The documents would have to be accompanied by XSL style sheets (one or more per document, each based on an underlying style for the template used).

As you rightly point out, most Word users couldn't care less about the structure of their documents or even the use of styles. However:

- the use of templates for letters, memos, etc. with suitable markup for header information could make them into searchable information resources at no cost to the user whatsoever;

- storing the document as valid XML means that it can be picked up and used by third-party software without any need for conversion or filtering;

- separating out markup from styles makes it much easier to support multiple views of the document within Word (online; paginated), and opens up the possibility of allowing users to fine-tune the styling of each view independently.

The second point is crucial to processes like publication, where a document will typically be authored in Word, and then converted at some point to a DTP package's internal format. Once that happens, any changes to the text will be made in the DTP environment, and not reflected in the 'source' Word document. The value of the out-of-date Word document is thus radically reduced, because it doesn't reflect the printed document correctly. In general terms, this limits the potential re-use of the document (even when all that is required is 'web page and DTP'ed hard copy').

If the Word document is actually XML, and the DTP package can work directly with XML documents and also save them as XML, then any changes made at the DTP stage will still be available from Word.

Storing Word documents as XML also opens up possibilities for search engines and databases to exploit the markup for more precise retrieval, as outlined in my first point.

In a way, you could see XML + XSL as the successor to RTF.

It seems to me that the biggest challenge to this vision that you face is getting documents well-formed and keeping them that way. (In our meeting we discussed the issue of horrible non-hierarchical HTML pages!)

Incidentally, I also have reservations about the use of HTML as any sort of storage format. It is fine as a delivery medium, and I like the way that XSL lets you deliver HTML - but that is its proper place in the overall scheme of things, in my view.

I could go into more detail (as you may gather, I have strong feelings on the subject!), but I thought it best to give you an initial set of comments and see what you think. Apart from anything else, I may have misunderstood the comments to which this email is a reaction!

I look forward to hearing from you, and getting your advice on how we might take this dialogue forward.

Best wishes,

Richard.

Richard Light

**MS/CR 0003127**  
**CONFIDENTIAL**

SGML/XML and Museum Information Consultancy  
richard@light.demon.co.uk

**MS/CR 0003128**  
**CONFIDENTIAL**