[MS-OINTXML]:

Office Intelligence Extensions to Office Open XML Structure

Intellectual Property Rights Notice for Open Specifications Documentation

- **Technical Documentation.** Microsoft publishes Open Specifications documentation ("this documentation") for protocols, file formats, data portability, computer languages, and standards support. Additionally, overview documents cover inter-protocol relationships and interactions.
- **Copyrights**. This documentation is covered by Microsoft copyrights. Regardless of any other terms that are contained in the terms of use for the Microsoft website that hosts this documentation, you can make copies of it in order to develop implementations of the technologies that are described in this documentation and can distribute portions of it in your implementations that use these technologies or in your documentation as necessary to properly document the implementation. You can also distribute in your implementation, with or without modification, any schemas, IDLs, or code samples that are included in the documentation. This permission also applies to any documents that are referenced in the Open Specifications documentation.
- **No Trade Secrets**. Microsoft does not claim any trade secret rights in this documentation.
- **Patents**. Microsoft has patents that might cover your implementations of the technologies described in the Open Specifications documentation. Neither this notice nor Microsoft's delivery of this documentation grants any licenses under those patents or any other Microsoft patents. However, a given Open Specifications document might be covered by the Microsoft Open Specifications Promise or the Microsoft Community Promise. If you would prefer a written license, or if the technologies described in this documentation are not covered by the Open Specifications Promise or Community Promise, as applicable, patent licenses are available by contacting iplg@microsoft.com.
- **License Programs**. To see all of the protocols in scope under a specific license program and the associated patents, visit the Patent Map.
- **Trademarks**. The names of companies and products contained in this documentation might be covered by trademarks or similar intellectual property rights. This notice does not grant any licenses under those rights. For a list of Microsoft trademarks, visit www.microsoft.com/trademarks.
- **Fictitious Names**. The example companies, organizations, products, domain names, email addresses, logos, people, places, and events that are depicted in this documentation are fictitious. No association with any real company, organization, product, domain name, email address, logo, person, place, or event is intended or should be inferred.

Reservation of Rights. All other rights are reserved, and this notice does not grant any rights other than as specifically described above, whether by implication, estoppel, or otherwise.

Tools. The Open Specifications documentation does not require the use of Microsoft programming tools or programming environments in order for you to develop an implementation. If you have access to Microsoft programming tools and environments, you are free to take advantage of them. Certain Open Specifications documents are intended for use in conjunction with publicly available standards specifications and network programming art and, as such, assume that the reader either is familiar with the aforementioned material or has immediate access to it.

Support. For questions and support, please contact dochelp@microsoft.com.

Revision Summary

Date	Revision History	Revision Class	Comments
4/22/2021	0.1	New	Released new document.
8/17/2021	1.0	Major	Significantly changed the technical content.

Table of Contents

1 Introduction

This document specifies elements and attributes for representing **observation** data, extending the XML vocabulary of the WordprocessingML file format described in [ISO/IEC29500-1:2016]. The new elements and attributes are presented using the extensibility mechanisms described in [ISO/IEC29500-3:2015]

Sections 1.7 and 2 of this specification are normative. All other sections and examples in this specification are informative.

1.1 Glossary

This document uses the following terms:

invalidation range: A range of text which specifies that some related item should be deleted if any part of the specified range is modified.

observation: A suggestion produced by a **workflow**.

workflow: An automated procedure that reads the open document and/or the user's actions and produces a set of suggested operations to assist with built-in or user-defined goals.

workflow type: An arbitrary Unicode name for a workflow which is not equal to any other known workflow type.

MAY, SHOULD, MUST, SHOULD NOT, MUST NOT: These terms (in all caps) are used as defined in [RFC2119]. All statements of optional behavior use either MAY, SHOULD, or SHOULD NOT.

1.2 References

Links to a document in the Microsoft Open Specifications library point to the correct section in the most recently published version of the referenced document. However, because individual documents in the library are not updated at the same time, the section numbers in the documents may not match. You can confirm the correct section numbering by checking the Errata.

1.2.1 Normative References

We conduct frequent surveys of the normative references to assure their continued availability. If you have any issue with finding a normative reference, please contact dochelp@microsoft.com. We will assist you in finding the relevant information.

[ISO/IEC29500-1:2016] ISO/IEC, "Information technology -- Document description and processing languages -- Office Open XML File Formats -- Part 1: Fundamentals and Markup Language Reference", ISO/IEC 29500-1:2016, https://www.iso.org/standard/71691.html

[ISO/IEC29500-2:2012] ISO/IEC, "Information technology -- Document description and processing languages -- Office Open XML File Formats -- Part 2: Open Packaging Conventions", ISO/IEC 29500-2:2012, http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=61796

[ISO/IEC29500-3:2015] ISO/IEC, "Information technology -- Document description and processing languages -- Office Open XML File Formats -- Part 3: Markup Compatibility and Extensibility", https://www.iso.org/standard/65533.html

[ISO/IEC29500-4:2016] ISO/IEC, "Information technology -- Document description and processing languages -- Office Open XML File Formats -- Part 4: Transitional Migration Features", https://www.iso.org/standard/71692.html

[MS-DOCX] Microsoft Corporation, "Word Extensions to the Office Open XML (.docx) File Format".

[MS-OEXTXML] Microsoft Corporation, "Office Shared Extensibility in Office Open XML".

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997, https://www.rfc-editor.org/rfc/rfc2119.html

[RFC3174] Eastlake III, D., and Jones, P., "US Secure Hash Algorithm 1 (SHA1)", RFC 3174, September 2001, http://www.ietf.org/rfc/rfc3174.txt

[RFC3629] Yergeau, F., "UTF-8, A Transformation Format of ISO 10646", STD 63, RFC 3629, November 2003, http://www.ietf.org/rfc/rfc3629.txt

[RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, October 2006, http://www.rfc-editor.org/rfc/rfc4648.txt

[URL] van Kesteren, A., "URL: Living Standard", June 2017, https://url.spec.whatwg.org/

[XMLSCHEMA1/2] Thompson, H., Beech, D., Maloney, M., and Mendelsohn, N., Eds., "XML Schema Part 1: Structures Second Edition", W3C Recommendation, October 2004, https://www.w3.org/TR/2004/REC-xmlschema-1-20041028/

[XMLSCHEMA2/2] Biron, P., and Malhotra, A., Eds., "XML Schema Part 2: Datatypes Second Edition", W3C Recommendation, October 2004, https://www.w3.org/TR/2004/REC-xmlschema-2-20041028/

1.2.2 Informative References

None.

1.3 Overview

This document specifies elements and attributes for representing data related to **observations**. This XML part enables arbitrary **workflows** to store data in a consistent way.

A workflows is a procedure that analyzes the document and/or user activity in order to provide assistance. Some possible examples of workflows include writing style analyzers, formatting assistants, and plagiarism recognizers. A workflow might be activated on request or might run automatically if they have been approved to run. This document describes the data that can be stored on the behalf of workflows; however, the operation of individual workflows is not generally described in this document. Each workflow is identified by a string known as the **workflow type**.

When a workflow runs, it provides zero or more observations. An observation is a suggested action. For example, a writing style analyzer might suggest replacing a specific word with a different word. Most observations are linked to a specific segment of the document, but some apply to the document as a whole.

Ordinarily, observation data is not arbitrarily deleted from the file even if the specified workflow type is not understood. However, many observations apply to a specific unit of text. These observations specify a range of text known as the **invalidation range**, which is a hint that the client should delete that observation if any part of the specified range is modified. For example, a grammar checker might set the observation's range to the exact location of a grammatical error and set the invalidation range to the entire surrounding sentence. This is because a change to any part of the sentence could potentially fix or change the grammatical error.

Of the data described in this document, the most frequently used portion is state information. For example, it is possible to store a record indicating that one specific observation has been rejected. Most state information is applicable even when nothing is known about the operation of that workflow.

State can be stored even with no additional details provided about an observation. For example, consider the case where a writing style analyzer produces an observation targeting a word of text, and this observation has then been rejected. The elements specified in this document can store a record of this rejection. When the same document is viewed again in the future, the writing style analyzer reaches the same conclusion and produces an identical observation again, but the observation can be suppressed because the stored state record exactly matches this observation.

Some specific workflows store data that is specific to their operation. This data is described in this document.

The elements and attributes specified in this document can also store information about which parts of the document have been processed by each workflow. This makes it possible to reduce resource utilization by only running the workflow against new content.

Finally, this document specifies how to store various settings that apply to the document.

1.4 Relationship to Protocols and Other Structures

This specification is dependent on the structures and concepts defined in the following references:

- [ISO/IEC29500-1:2016] for the baseline WordprocessingML persistence format.
- [ISO/IEC29500-2:2012] for open packaging conventions.
- [ISO/IEC29500-3:2015] for markup compatibility and extensibility.
- [ISO/IEC29500-4:2016] for backwards compatibility considerations.
- [MS-DOCX] for WordprocessingML extensions.
- [MS-OEXTXML] for complex types for extension lists.
- [RFC3174] for the SHA-1 hash code algorithm.
- [RFC4648] for Base64 format.

1.5 Applicability Statement

This document specifies a persistence format for extensions as specified by [ISO/IEC29500-1:2016] for WordprocessingML documents. The extensions specified in this document enable expressing **observation** data and some related metadata and are not applicable as a stand-alone file format. Each structure specified in this document is integrated with WordprocessingML documents in a particular way as specified in the section for that structure. All structures are integrated into WordprocessingML documents in a way that maintains compatibility with [ISO/IEC29500-1:2016] implementations. This persistence format can also be used for files that do not use WordprocessingML as long as no elements referencing WordprocessingML are used.

The extensions specified in this document do not require any other extensions to be used and do not prohibit any other extensions from being used in the same document.

1.6 Versioning and Localization

Certain XML elements specified in this document can specify a version number. The client should ignore any element which specifies a version other than what the client expects.

1.7 Vendor-Extensible Fields

Vendors MAY specify any desired value for the **workflow type** of a CT_State (section 2.1.3.7) or CT_OnDemandWorkflow (section 2.1.3.14) element.

Vendors MAY specify any desired state value within a CT_State (section 2.1.3.7) element.

2 Structures

2.1 http://schemas.microsoft.com/office/intelligence/2020/intelligence

2.1.1 Elements

2.1.1.1 intelligence

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

A CT_Intelligence element that specifies the root element for the entire XML part.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this element.

```
<xsd:element name="intelligence" type="CT_Intelligence"/>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.1.2 goals

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

A CT_Goals element that specifies a setting describing the author's objective. The setting can be used to influence **observations** that are shown. This element MUST NOT be used as the root of the XML part. One instance of this element MAY be used in each instance of oel:CT_Extension ([MS-OEXTXML] section 2.1.3.1) specifying the URI "74B372B9-2EFF-4315-9A3F-32BA87CA82B1" within an instance of oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) within an instance of CT_IntelligenceSettings (section 2.1.3.15).

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this element.

```
<xsd:element name="goals" type="CT_Goals"/>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.1.3 similarityCritique

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

A CT_SimilarityCritique element that specifies data for an **observation** indicating that a portion of the document appears similar to another source. This element MUST NOT be used as the root of the XML part. One instance of this element MAY be used in each instance of oel:CT_Extension ([MS-OEXTXML] section 2.1.3.1) specifying the URI "426473B9-03D8-482F-96C9-C2C85392BACA" within an instance of oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) within an instance of CT_Content (section 2.1.3.3).

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this element.

```
<xsd:element name="similarityCritique" type="CT_SimilarityCritique"/>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.1.4 similaritySummary

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

A CT_SimilaritySummary element that specifies a record of an **observation** which reports statistics related to CT_SimilarityCritique (section 2.1.3.8). This element MUST NOT be used as the root of the XML part. One instance of this element MAY be used in each instance of oel:CT_Extension ([MS-OEXTXML] section 2.1.3.1) specifying the URI "E302BA01-7950-474C-9AD3-286E660C40A8" within an instance of oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) within an instance of CT_Content (section 2.1.3.3).

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this element.

```
<xsd:element name="similaritySummary" type="CT_SimilaritySummary"/>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.2 Attributes

None.

2.1.3 Complex Types

2.1.3.1 CT_Intelligence

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: intelligence

Specifies the root element for the entire XML part.

Child Elements:

observations: A CT_Observations element that specifies the root element for all **observation** data in the XML part.

intelligenceSettings: A CT_IntelligenceSettings element that specifies the root element for settings.

onDemandWorkflows: A CT_OnDemandWorkflows element that specifies the root element for workflow progress data.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_Intelligence">
  <xsd:sequence>
    <xsd:element name="observations" type="CT_Observations" minOccurs="0" maxOccurs="1"/>
    <xsd:element name="intelligenceSettings" type="CT_IntelligenceSettings" minOccurs="0"
maxOccurs="1"/>
    <xsd:element name="onDemandWorkflows" type="CT OnDemandWorkflows" minOccurs="0"
maxOccurs="1"/>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
```

```
  </xsd:sequence>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.2 CT_Observations

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_Intelligence

Specifies the root element for all **observation** data in the XML part.

Child Elements:

textHash: A CT_TextHash (section 2.1.3.4) element that specifies a reference to all runs of text in the document which match a specified hash code.

bookmark: A CT_Bookmark (section 2.1.3.5) element that specifies a reference to text within a bookmark ([ISO/IEC29500-1:2016] section 17.3.6) in the document.

entireDocument: A CT_EntireDocument (section 2.1.3.6) element that specifies a reference to the entire document as a whole.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_Observations">
  <xsd:sequence>
    <xsd:element name="textHash" type="CT_TextHash" minOccurs="0" maxOccurs="unbounded"/>
    <xsd:element name="bookmark" type="CT_Bookmark" minOccurs="0" maxOccurs="unbounded"/>
    <xsd:element name="entireDocument" type="CT_EntireDocument" minOccurs="0"
maxOccurs="unbounded"/>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.3 CT_Content

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_TextHash, CT_Bookmark, CT_EntireDocument

The base type for CT_TextHash (section 2.1.3.4), CT_Bookmark (section 2.1.3.5), and CT_EntireDocument (section 2.1.3.6). Specifies a reference to one or more specific units of content in the document.

Child Elements:

state: A CT_State (section 2.1.3.7) element that specifies the current state of an **observation** that affects the referenced unit of content.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies current and future extensions.

Attributes:

id: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies a unique identifier for this element. The specified value is assumed to be unique among all sibling elements. If more than one element specifies the same ID value, only the first matching element SHOULD be used.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_Content">
  <xsd:sequence>
    <xsd:element name="state" type="CT State" minOccurs="0" maxOccurs="unbounded"/>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="id" type="xsd:string" use="required"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.4 CT_TextHash

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_Observations

Specifies a reference to all runs of text in the document which match a specified hash code. This can simultaneously refer to multiple different locations if the same text appears in each location. Matching is case-sensitive because of the way the hash code is computed.

Child Elements:

state: A CT_State (section 2.1.3.7) element that specifies the current state of an **observation** that affects the referenced unit of content.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies current and future extensions.

Attributes:

hashCode: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the hash code of any amount of text. This MUST be determined by representing the text as a UTF-8 octet stream ([RFC3629]), computing the SHA-1 ([RFC3174]) hash code of the stream, representing the SHA-1 ([RFC3174]) hash code in Base64 ([RFC4648]) format, and finally taking only the first 14 characters of the result.

id: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies a unique identifier for this element. The specified value is assumed to be unique among all sibling elements. If more than one element specifies the same ID value, only the first matching element SHOULD be used.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_TextHash">
  <xsd:complexContent>
    <xsd:extension base="CT_Content">
      <xsd:attribute name="hashCode" type="xsd:string" use="required"/>
    </xsd:extension>
```

```
  </xsd:complexContent>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.5 CT_Bookmark

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_Observations

Specifies a reference to text within a specified bookmark ([ISO/IEC29500-1:2016] section 17.3.6) in the document.

Child Elements:

state: A CT_State (section 2.1.3.7) element that specifies the current state of an **observation** that affects the referenced unit of content.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies current and future extensions.

Attributes:

bookmarkName: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the name of the bookmark ([ISO/IEC29500-1:2016] section 17.3.6) showing the location of the item being referenced. The value MUST begin with "_Int_". This element SHOULD be ignored if the specified bookmark ([ISO/IEC29500-1:2016] section 17.3.6) has its w:bookmarkStart ([ISO/IEC29500-1:2016] section 17.3.6.2) and w:bookmarkEnd ([ISO/IEC29500-1:2016] section 17.3.6.1) tags in two different paragraphs.

invalidationBookmarkName: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that optionally specifies the name of the bookmark ([ISO/IEC29500-1:2016] section 17.3.6) showing the **invalidation range**. The value MUST begin with "_Int_". This value hints that the element can be deleted if any modification is applied to the text within the invalidation range. If this attribute is not set, the bookmark ([ISO/IEC29500-1:2016] section 17.3.6) referenced by bookmarkName constitutes the invalidation range by default. This element SHOULD be ignored if this specified bookmark ([ISO/IEC29500-1:2016] section 17.3.6) has start and end markers in two different paragraphs, or if it exists on a different paragraph than the bookmark ([ISO/IEC29500-1:2016] section 17.3.6) specified by bookmarkName.

hashCode: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that optionally specifies the hash code of the text in the invalidation range. If this attribute is present, the value MUST be determined by representing the text as a UTF-8 octet stream ([RFC3629]), computing the SHA-1 ([RFC3174]) hash code of the stream, representing the SHA-1 ([RFC3174]) hash code in Base64 ([RFC4648]) format, and finally taking only the first 14 characters of the result.

id: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies a unique identifier for this element. The specified value is assumed to be unique among all sibling elements. If more than one element specifies the same ID value, only the first matching element SHOULD be used.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_Bookmark">
  <xsd:complexContent>
    <xsd:extension base="CT_Content">
      <xsd:attribute name="bookmarkName" type="xsd:string" use="required"/>
      <xsd:attribute name="invalidationBookmarkName" type="xsd:string" use="optional"/>
```

```
      <xsd:attribute name="hashCode" type="xsd:string" use="optional"/>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.6 CT_EntireDocument

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_Observations

Specifies a reference to the entire document as a whole.

Child Elements:

state: A CT_State (section 2.1.3.7) element that specifies the current state of an **observation** that affects the referenced unit of content.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies current and future extensions.

Attributes:

id: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies a unique identifier for this element. The specified value is assumed to be unique among all sibling elements. If more than one element specifies the same ID value, only the first matching element SHOULD be used.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_EntireDocument">
  <xsd:complexContent>
    <xsd:extension base="CT_Content"/>
  </xsd:complexContent>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.7 CT_State

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_TextHash, CT_Content, CT_Bookmark, CT_EntireDocument

Specifies the current state of an **observation** affecting the unit of content described by the parent element.

Child Elements:

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

type: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the **workflow type** which this state applies to.

value: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the state of the observation. The value "Rejected" indicates that the observation was expressly rejected. The value "Reviewed" indicates that the observation was manually reviewed in some way. No other values currently have any effect, but this might be extended in the future.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_State">
  <xsd:sequence>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="type" type="xsd:string" use="required"/>
  <xsd:attribute name="value" type="xsd:string" use="required"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.8 CT_SimilarityCritique

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: similarityCritique

Specifies data constituting an **observation** indicating that a portion of the document appears similar to another source.

Child Elements:

source: A CT_SimilaritySource (section 2.1.3.9) element that specifies a similar source of content**.**

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

version: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that SHOULD equal 1, otherwise behavior is unspecified.

context: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies an arbitrary selection of text from the vicinity of the affected range of text. This is used to display previews of the suggested operations.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_SimilarityCritique">
  <xsd:sequence>
    <xsd:element name="source" type="CT SimilaritySource" minOccurs="0"
maxOccurs="unbounded"/>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="version" type="xsd:int"/>
  <xsd:attribute name="context" type="xsd:string"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.9 CT_SimilaritySource

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_SimilarityCritique

Specifies a source which contains content similar to what is found in the document.

Child Elements:

suggestions: A CT_SimilaritySuggestionsForType (section 2.1.3.10) element that specifies the collection of all suggested rewrites for the apparently similar text that apply for a given citation type.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

sourceType: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that SHOULD equal "Online", otherwise behavior is unspecified.

sourceTitle: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the title of the similar source. For example, this could be the title of a web page where similar text was found.

sourceUrl: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies a URL ([URL]) identifying to the similar source.

sourceSnippet: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies a selection of matching text from the similar source. This is used to display a preview of the suggested operation.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_SimilaritySource">
  <xsd:sequence>
    <xsd:element name="suggestions" type="CT_SimilaritySuggestionsForType" minOccurs="0"
maxOccurs="unbounded"/>
    <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="sourceType" type="xsd:string"/>
  <xsd:attribute name="sourceTitle" type="xsd:string"/>
  <xsd:attribute name="sourceUrl" type="xsd:string"/>
  <xsd:attribute name="sourceSnippet" type="xsd:string"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.10 CT_SimilaritySuggestionsForType

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_SimilaritySource

Specifies the collection of all suggested rewrites for the apparently similar text that apply for a given citation type.

Child Elements:

suggestion: A CT_SimilaritySuggestion (section 2.1.3.11) element that specifies a single suggested citation. Each instance of CT_SimilaritySuggestionsForType SHOULD have three of these children, specifying each of the following values of citationStyle: “Mla”, “Apa”, and “Chicago”.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

citationType: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the type of citation. Supported values are "Full", indicating that a citation is added as a block, or "Inline", indicating that a citation is typically appended to the text. Other values currently have no effect, but this might be extended in the future.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_SimilaritySuggestionsForType">
  <xsd:sequence>
    <xsd:element name="suggestion" type="CT SimilaritySuggestion" minOccurs="0"
maxOccurs="unbounded"/>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="citationType" type="xsd:string"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.11 CT_SimilaritySuggestion

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_SimilaritySuggestionsForType

Specifies a single suggested citation.

Child Elements:

citationText: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) element that specifies suggested text that can be appended to the matching text range as a citation.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

citationStyle: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the style guide that the citation is based on. Supported values are "Mla", "Apa", and "Chicago". Other values currently have no effect, but this might be extended in the future.

isIdentical: A xsd:boolean ([XMLSCHEMA2/2] section 3.2.2) attribute that specifies whether the text in the document exactly matches the other source.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_SimilaritySuggestion">
  <xsd:sequence>
    <xsd:element name="citationText" type="xsd:string"/>
    <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="citationStyle" type="xsd:string"/>
  <xsd:attribute name="isIdentical" type="xsd:boolean"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.12 CT_SimilaritySummary

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: similaritySummary

Specifies data for an **observation** which reports statistics related to CT_SimilarityCritique (section 2.1.3.8).

Child Elements:

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

version: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that SHOULD equal 1, otherwise behavior is unspecified.

runId: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies an arbitrary ID representing the origin of the other attribute values applied to this element. This value is not useful when loading a document for the first time in a session. This value is useful when CT_SimilarityCritique (section 2.1.3.8) elements are created by some **workflow** that is not finished running by the time the file is saved. In this case, if any more observations are produced, the runId value can be used to determine whether those observations will affect the other attributes of this element, or alternatively whether most values will restart from zero.

tilesCheckedInThisRun: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that specifies the number of paragraphs that have been checked for similarity to other sources.

totalNumOfTiles: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that specifies the total number of paragraphs in the document that could potentially be checked for similarity.

similarityAnnotationCount: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that specifies the number of SimilarityCritique observations that are currently in the document.

numWords: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that specifies the total number of words in the document that have been checked for similarity to other sources.

numFlaggedWords: A xsd:int ([XMLSCHEMA2/2] section 3.3.17) attribute that specifies the total number of words in the document that are considered similar to other sources.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_SimilaritySummary">
  <xsd:sequence>
    <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="version" type="xsd:int"/>
  <xsd:attribute name="runId" type="xsd:string"/>
  <xsd:attribute name="tilesCheckedInThisRun" type="xsd:int"/>
  <xsd:attribute name="totalNumOfTiles" type="xsd:int"/>
  <xsd:attribute name="similarityAnnotationCount" type="xsd:int"/>
  <xsd:attribute name="numWords" type="xsd:int"/>
  <xsd:attribute name="numFlaggedWords" type="xsd:int"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.13 CT_OnDemandWorkflows

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_Intelligence

Specifies the root element for **workflow** progress data.

Child Elements:

onDemandWorkflow: A CT_OnDemandWorkflow element that specifies the current level of completeness of a workflow which processes paragraphs.

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_OnDemandWorkflows">
  <xsd:sequence>
    <xsd:element name="onDemandWorkflow" type="CT_OnDemandWorkflow" minOccurs="0"
maxOccurs="unbounded"/>
    <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.14 CT_OnDemandWorkflow

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_OnDemandWorkflows

Specifies the current level of completeness of a **workflow** which processes paragraphs.

Child Elements:

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

type: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the **workflow type** which this record applies to.

paragraphVersions: An ST_ParagraphVersions attribute that specifies a space-separated list of paragraphs which the specified workflow has finished processing. Each paragraph is specified in the format "{paraId}-{textId}" ([MS-DOCX] section 2.6.2.3) ([MS-DOCX] section 2.6.2.4). For example, a possible value is "11111111-AAAAAAAA 22222222-BBBBBBBB 01234567-89ABCDEF".

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_OnDemandWorkflow">
  <xsd:sequence>
```

```
    <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="type" type="xsd:string" use="required"/>
  <xsd:attribute name="paragraphVersions" type="ST_ParagraphVersions" use="required"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.15 CT_IntelligenceSettings

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_Intelligence

Specifies the root element for settings.

Child Elements:

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies current and future extensions.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_IntelligenceSettings">
  <xsd:sequence>
    <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.3.16 CT_Goals

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: goals

Specifies a setting describing the author's objective. The setting can be used to influence **observations** that are shown.

Child Elements:

extLst: An oel:CT_ExtensionList ([MS-OEXTXML] section 2.1.3.2) element that specifies future extensions.

Attributes:

version: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that SHOULD equal 1, otherwise behavior is unspecified.

formality: A xsd:string ([XMLSCHEMA2/2] section 3.2.1) attribute that specifies the quantity of observations that would be displayed. The value "0" indicates that all available observations SHOULD be shown. The value "1" indicates that some less-important observations SHOULD be hidden from view. The value "2" indicates that only the most important observations SHOULD be displayed. No other values are supported.

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this complex type.

```
<xsd:complexType name="CT_Goals">
  <xsd:sequence>
    <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
  </xsd:sequence>
  <xsd:attribute name="version" type="xsd:string"/>
  <xsd:attribute name="formality" type="xsd:string"/>
</xsd:complexType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

2.1.4 Simple Types

2.1.4.1 ST_ParagraphVersions

Target namespace: http://schemas.microsoft.com/office/intelligence/2020/intelligence

Referenced by: CT_OnDemandWorkflow

Specifies a space-separated list of paragraphs which the specified workflow has finished processing. Each paragraph is specified in the format "{paraId}-{textId}" ([MS-DOCX] section 2.6.2.3) ([MS-DOCX] section 2.6.2.4). For example, a possible value is "11111111-AAAAAAAA 22222222-BBBBBBBB 01234567-89ABCDEF".

The following W3C XML Schema ([XMLSCHEMA1/2] section 2.1) fragment specifies the contents of this simple type.

```
<xsd:simpleType name="ST_ParagraphVersions">
  <xsd:list itemType="xsd:string"/>
</xsd:simpleType>
```

See section 5.1 for the full W3C XML Schema ([XMLSCHEMA1/2] section 2.1).

3 Structure Examples

3.1 Ignore All

This example shows how to use the types in this document to specify that a certain kind of observation should never be displayed if it would affect an exact word or phrase.

"WritingAssistant" is a hypothetical **workflow type** for a **workflow** that produces **observations** related to the writing style of a document. This workflow produces an observation every time the word "whom" is used in a document. The user reviews one such observation and decides that it should never be displayed again even if the word "whom" is used again in another location.

The content selector CT_TextHash (section 2.1.3.4) can be used to suppress all such instances. This element requires a hash code. The SHA-1 hash code ([RFC3174]) of the word "whom" represented in Base64 format ([RFC4648]) is "CXaroNQwQFYioAmBvFjG8W0qQPE=". The first 14 characters, "CXaroNQwQFYioA", are the final hash code.

Any arbitrary string is permitted for the value of *id* as long as the value is unique among all sibling elements.

```
<int2:intelligence
xmlns:int2="http://schemas.microsoft.com/office/intelligence/2020/intelligence"
xmlns:oel="http://schemas.microsoft.com/office/2019/extlst">
  <int2:observations>
    <int2:textHash int2:hashCode="CXaroNQwQFYioA" int2:id="abc">
      <int2:state int2:type="WritingAssistant" int2:value="Rejected"/>
    </int2:textHash>
  </int2:observations>
</int2:intelligence>
```

3.2 Review Specific Observation

This example shows how to use the types in this document to specify that a specific instance of an **observation** has been reviewed by the user.

This example uses the same hypothetical **workflow type** as the previous example (section 3.1), "WritingAssistant". The user notices one observation applied to the word "whom" and chooses to indicate that the observation has been reviewed even though the suggestion was not applied to the document.

The content selector CT_Bookmark (section 2.1.3.5) can be used to save this state. First, a bookmark ([ISO/IEC29500-1:2016] section 17.3.6) must be created around the particular instance of the word "whom", and the bookmark name must begin with "_Int_". For this example, the selected bookmark name is "_Int_12345". The observation does not depend on any other words in the sentence to remain applicable, so it is appropriate to leave the invalidation range attribute unset. However, it is strongly encouraged to provide the hash code of the invalidation range, which is "CXaroNQwQFYioA" (see the previous example, section 3.1).

Any arbitrary string is permitted for the value of *id* as long as the value is unique among all sibling elements.

```
<int2:intelligence
xmlns:int="http://schemas.microsoft.com/office/intelligence/2020/intelligence"
xmlns:oel="http://schemas.microsoft.com/office/2019/extlst">
  <int2:observations>
    <int2:bookmark int2:bookmarkName="_Int_12345" int2:hashCode="CXaroNQwQFYioA"
int2:id="abc">
      <int2:state int2:type="WritingAssistant" int2:value="Reviewed" />
```

```
    </int2:bookmark >
  </int2:observations>
</int2:intelligence>
```

3.3 Review Specific Observation With Invalidation Range

This example demonstrates when it might be appropriate to set the invalidationBookmarkName attribute on instances of CT_Bookmark (section 2.1.3.5).

In this example, a hypothetical **workflow type** "GrammarChecker" produces an **observation** for the following paragraph. The observation applies to the **bold** text, and the invalidation range is underlined.

Lorem ipsum dolar sit amet. The quick brown fox **jump** over the lazy dog. Lorem ipsum dolar sit amet.

Logically, the reason this observation was produced is that "jump" should be changed to "jumped". However, this is not true in general; the word "jump" may be used without issue in many other possible sentences. This means that the observation may be rendered invalid if any portion of the sentence is modified. However, this does not extend to other sentences in the paragraph.

First, a bookmark ([ISO/IEC29500-1:2016] section 17.3.6) must be created around the word "jump" and the bookmark name must begin with "_Int_". For this example, the selected bookmark name is "_Int_12345". Next, a second bookmark ([ISO/IEC29500-1:2016] section 17.3.6) must be created around the sentence "The quick brown fox jump over the lazy dog." The second bookmark's name must also begin with "_Int_". For this example, the second bookmark's name is "_Int_67890". Finally, a hash code should be computed for the sentence "The quick brown fox jump over the lazy dog." The computed hash code is "PCRd4lSIsx4R/A" (see section 3.1).

```
<int2:intelligence
xmlns:int="http://schemas.microsoft.com/office/intelligence/2020/intelligence"
xmlns:oel="http://schemas.microsoft.com/office/2019/extlst">
  <int2:observations>
    <int2:bookmark int2:bookmarkName="_Int_12345" int2:invalidationBookmarkName="_Int_67890"
int2:hashCode="PCRd4lSIsx4R/A" int2:id="abc">
      <int2:state int2:type="GrammarChecker" int2:value="Reviewed" />
    </int2:bookmark >
  </int2:observations>
</int2:intelligence>
```

3.4 Workflow progress

This example shows a situation when CT_OnDemandWorkflows (section 2.1.3.13) may be useful.

In this example, a hypothetical **workflow type** "DocumentProcessor" scans the document one paragraph at a time and does unspecified work. This scanning process takes a very long time, and it is very likely that the file will be saved while the scan is still in progress. If the scan restarts on each save, it might never have an opportunity to finish.

The CT_OnDemandWorkflow (section 2.1.3.14) element can record which paragraphs have been scanned. Each paragraph is listed by its paraId ([MS-DOCX] section 2.6.2.3) and textId ([MS-DOCX] section 2.6.2.4). The textId ([MS-DOCX] section 2.6.2.4) value ensures that the scanning of a paragraph will be forgotten if that paragraph is modified.

This sample XML shows that the **workflow** has processed three paragraphs. There may or may not be more paragraphs that still need to be processed.

```
<int2:intelligence
xmlns:int="http://schemas.microsoft.com/office/intelligence/2020/intelligence"
xmlns:oel="http://schemas.microsoft.com/office/2019/extlst">
  <int2:onDemandWorkflows>
    <int2:onDemandWorkflow int2:type="DocumentProcessor" int2:paragraphVersions="11111111-
AAAAAAAA 22222222-BBBBBBBB 33333333-CCCCCCCC" />
  </int2:onDemandWorkflows>
</int2:intelligence>
```

3.5 Goals setting

This example shows how to use the formality goal setting to control how many **observations** should be displayed.

The user decides that the least important observations should not be displayed. This decision can be stored with the following XML.

```
<int2:intelligence
xmlns:int="http://schemas.microsoft.com/office/intelligence/2020/intelligence"
xmlns:oel="http://schemas.microsoft.com/office/2019/extlst">
  <int2:intelligenceSettings>
    <int2:extLst>
      <oel:ext oel:uri="74B372B9-2EFF-4315-9A3F-32BA87CA82B1">
        <int2:goals int2:version="1" int2:formality="1" />
      </oel:ext>
    </int2:extLst>
  </int2:intelligenceSettings>
</int2:intelligence>
```

4 Security

4.1 Security Considerations for Implementers

Because hashing may be reversible, hash codes stored in this XML part should be assigned the same level of protection as the content that was hashed.

4.2 Index of Security Fields

None.

5 Appendix A: Full XML Schemas

Schema name	Prefix	Section
http://schemas.microsoft.com/office/intelligence/2020/intelligence Schema	None.	5.1

5.1 http://schemas.microsoft.com/office/intelligence/2020/intelligence Schema

```
<xsd:schema xmlns:oel="http://schemas.microsoft.com/office/2019/extlst"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:w12="http://schemas.openxmlformats.org/wordprocessingml/2006/main"
elementFormDefault="qualified" blockDefault="#all"
xmlns="http://schemas.microsoft.com/office/intelligence/2020/intelligence"
targetNamespace="http://schemas.microsoft.com/office/intelligence/2020/intelligence">
  <xsd:import id="oel" namespace="http://schemas.microsoft.com/office/2019/extlst"
schemaLocation="officeextlst.xsd"/>
  <xsd:element name="intelligence" type="CT_Intelligence"/>
  <xsd:complexType name="CT Intelligence">
    <xsd:sequence>
      <xsd:element name="observations" type="CT_Observations" minOccurs="0" maxOccurs="1"/>
      <xsd:element name="intelligenceSettings" type="CT_IntelligenceSettings" minOccurs="0"
maxOccurs="1"/>
      <xsd:element name="onDemandWorkflows" type="CT_OnDemandWorkflows" minOccurs="0"
maxOccurs="1"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="CT_Observations">
    <xsd:sequence>
      <xsd:element name="textHash" type="CT TextHash" minOccurs="0" maxOccurs="unbounded"/>
      <xsd:element name="bookmark" type="CT_Bookmark" minOccurs="0" maxOccurs="unbounded"/>
      <xsd:element name="entireDocument" type="CT EntireDocument" minOccurs="0"
maxOccurs="unbounded"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="CT_Content">
    <xsd:sequence>
      <xsd:element name="state" type="CT_State" minOccurs="0" maxOccurs="unbounded"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="id" type="xsd:string" use="required"/>
  </xsd:complexType>
  <xsd:complexType name="CT_TextHash">
    <xsd:complexContent>
      <xsd:extension base="CT_Content">
        <xsd:attribute name="hashCode" type="xsd:string" use="required"/>
      </xsd:extension>
    </xsd:complexContent>
  </xsd:complexType>
  <xsd:complexType name="CT_Bookmark">
    <xsd:complexContent>
      <xsd:extension base="CT Content">
        <xsd:attribute name="bookmarkName" type="xsd:string" use="required"/>
        <xsd:attribute name="invalidationBookmarkName" type="xsd:string" use="optional"/>
        <xsd:attribute name="hashCode" type="xsd:string" use="optional"/>
      </xsd:extension>
    </xsd:complexContent>
  </xsd:complexType>
  <xsd:complexType name="CT_EntireDocument">
    <xsd:complexContent>
      <xsd:extension base="CT_Content"/>
    </xsd:complexContent>
  </xsd:complexType>
```

```
  <xsd:complexType name="CT State">
    <xsd:sequence>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="type" type="xsd:string" use="required"/>
    <xsd:attribute name="value" type="xsd:string" use="required"/>
  </xsd:complexType>
  <xsd:complexType name="CT_IntelligenceSettings">
    <xsd:sequence>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="CT_OnDemandWorkflows">
    <xsd:sequence>
      <xsd:element name="onDemandWorkflow" type="CT_OnDemandWorkflow" minOccurs="0"
maxOccurs="unbounded"/>
      <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
  </xsd:complexType>
  <xsd:complexType name="CT_OnDemandWorkflow">
    <xsd:sequence>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="type" type="xsd:string" use="required"/>
    <xsd:attribute name="paragraphVersions" type="ST_ParagraphVersions" use="required"/>
  </xsd:complexType>
  <xsd:simpleType name="ST_ParagraphVersions">
    <xsd:list itemType="xsd:string"/>
  </xsd:simpleType>
  <xsd:element name="similarityCritique" type="CT_SimilarityCritique"/>
  <xsd:complexType name="CT SimilarityCritique">
    <xsd:sequence>
      <xsd:element name="source" type="CT_SimilaritySource" minOccurs="0"
maxOccurs="unbounded"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="version" type="xsd:int"/>
    <xsd:attribute name="context" type="xsd:string"/>
  </xsd:complexType>
  <xsd:complexType name="CT SimilaritySource">
    <xsd:sequence>
      <xsd:element name="suggestions" type="CT SimilaritySuggestionsForType" minOccurs="0"
maxOccurs="unbounded"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="sourceType" type="xsd:string"/>
    <xsd:attribute name="sourceTitle" type="xsd:string"/>
    <xsd:attribute name="sourceUrl" type="xsd:string"/>
    <xsd:attribute name="sourceSnippet" type="xsd:string"/>
  </xsd:complexType>
  <xsd:complexType name="CT_SimilaritySuggestionsForType">
    <xsd:sequence>
      <xsd:element name="suggestion" type="CT_SimilaritySuggestion" minOccurs="0"
maxOccurs="unbounded"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="citationType" type="xsd:string"/>
  </xsd:complexType>
  <xsd:complexType name="CT SimilaritySuggestion">
    <xsd:sequence>
      <xsd:element name="citationText" type="xsd:string"/>
      <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    <xsd:attribute name="citationStyle" type="xsd:string"/>
    <xsd:attribute name="isIdentical" type="xsd:boolean"/>
  </xsd:complexType>
  <xsd:element name="similaritySummary" type="CT_SimilaritySummary"/>
  <xsd:complexType name="CT_SimilaritySummary">
    <xsd:sequence>
```

```
        <xsd:element name="extLst" type="oel:CT ExtensionList" minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
      <xsd:attribute name="version" type="xsd:int"/>
      <xsd:attribute name="runId" type="xsd:string"/>
      <xsd:attribute name="tilesCheckedInThisRun" type="xsd:int"/>
      <xsd:attribute name="totalNumOfTiles" type="xsd:int"/>
      <xsd:attribute name="similarityAnnotationCount" type="xsd:int"/>
      <xsd:attribute name="numWords" type="xsd:int"/>
      <xsd:attribute name="numFlaggedWords" type="xsd:int"/>
    </xsd:complexType>
    <xsd:element name="goals" type="CT Goals"/>
    <xsd:complexType name="CT_Goals">
      <xsd:sequence>
        <xsd:element name="extLst" type="oel:CT_ExtensionList" minOccurs="0" maxOccurs="1"/>
      </xsd:sequence>
      <xsd:attribute name="version" type="xsd:string"/>
      <xsd:attribute name="formality" type="xsd:string"/>
    </xsd:complexType>
  </xsd:schema>
```

6 Appendix B: Product Behavior

The information in this specification is applicable to the following Microsoft products or supplemental software. References to product versions include updates to those products.

- Microsoft Word Online
- Microsoft Word 2021

Exceptions, if any, are noted in this section. If an update version, service pack or Knowledge Base (KB) number appears with a product name, the behavior changed in that update. The new behavior also applies to subsequent updates unless otherwise specified. If a product edition appears with the product version, behavior is different in that product edition.

Unless otherwise specified, any statement of optional behavior in this specification that is prescribed using the terms "SHOULD" or "SHOULD NOT" implies product behavior in accordance with the SHOULD or SHOULD NOT prescription. Unless otherwise specified, the term "MAY" implies that the product does not follow the prescription.

7 Change Tracking

This section identifies changes that were made to this document since the last release. Changes are classified as Major, Minor, or None.

The revision class **Major** means that the technical content in the document was significantly revised. Major changes affect protocol interoperability or implementation. Examples of major changes are:

- A document revision that incorporates changes to interoperability requirements.
- A document revision that captures changes to protocol functionality.

The revision class **Minor** means that the meaning of the technical content was clarified. Minor changes do not affect protocol interoperability or implementation. Examples of minor changes are updates to clarify ambiguity at the sentence, paragraph, or table level.

The revision class **None** means that no new technical changes were introduced. Minor editorial and formatting changes may have been made, but the relevant technical content is identical to the last released version.

The changes made to this document are listed in the following table. For more information, please contact dochelp@microsoft.com.

Section	Description	Revision class
2.1.1.2 goals	Updated link to CT_ExtensionList.	Minor
2.1.1.3 similarityCritique	Updated link to CT_ExtensionList.	Minor
2.1.3.12 CT_SimilaritySummary	Changed should to will.	Minor
2.1.3.16 CT_Goals	Changed should to would.	Minor

8 Index