

The Millistream News Format

Technical Description and Reference Guide

5 May 2021

Introduction

The Millistream News Format is a very scaled down version of the NewsML and NITF formats. Due to the heritage, a transition to the full NewsML/NITF formats is quite possible if customer demand for those formats would show. Data normally stored inside the document itself in NewsML has been placed in the *MDF_M_NEWSCONTENT* message (users of the MDA will find this data in the 'news' and 'news coding' tables) and thus removed from this specification due to redundancy.

The News Format

The top XML element in a Millistream News document, *NewsItem*, looks like this:

```
<?xml version="1.0" encoding="UTF-8" ?>
<NewsItem>
  <NewsEnvelope>
    [...]
  </NewsEnvelope>
  <ContentItem>
    [...]
  </ContentItem>
</NewsItem>
```

The first line indicates that this file complies with the XML version 1.0 specification and that the character encoding used throughout the file is UTF-8, which is the only supported character encoding within the Millistream system.

There are two required markup components:

- a required, non-repeatable *NewsEnvelope* element, containing information about this specific transmission
- one or more required, repeatable *ContentItem* elements, containing the main content

News Envelope

The *NewsEnvelope* element holds transmission information, the equivalent of the writing on the outside of a virtual envelope holding the news story, and typically looks like this:

```
<NewsEnvelope>
  <TransmissionId>20090227:BIT:2283:0</TransmissionId>
  <DateAndTime>20090302T090000+0100</DateAndTime>
  <SentFrom>
    <Party FormalName="Ticket Travel Group" />
  </SentFrom>
</NewsEnvelope>
```

There are three required markup components:

- a required, non-repeatable *TransmissionId* element, containing a specific unique identifier for the transmission as sent by the source. Since this value is set by the source, Millistream can make no guarantees regarding the uniqueness of this value, please use the *MDF_F_NEWSID* field (the 'news.news id' column in the MDA) to get a truly unique id of this news item.
- A required, non-repeatable *DateAndTime* element, containing the date and time of the transmission in ISO 8601 Basic format.
- A required, non-repeatable *SentFrom* element, containing a *Party* subelement identifying the sender using a formal name. For News Agencies this will be the name of the News Agency, and for Press Release Agencies this will be the name of the company issuing the Press Release.

Content Items

The *ContentItem* either contains the actual resource or holds a pointer to it. Textual resources, like news stories (including XML files), will usually be included internally; binary resources, like photos or PDF documents, must be included externally. Here is an example of a content item referencing an external file holding a PDF document:

```
<ContentItem Href="http://millistream.com/files/test.pdf">  
  <MimeType FormalName="application/pdf" />  
</ContentItem>
```

The *ContentItem* element can contain up to two of the following three markup components:

- an optional *Href* attribute containing the relative or absolute URL of the content file
- an optional *Role* attribute defining the role of this content item. Current valid values:
 - “*FactBox*” - The content item contains a Fact Box type of text.
 - “*External*” - The content item links to an external document.
- a required *MimeType* element containing the MIME type of the file
- an optional *DataContent* element containing the actual textual or XML content.

The *Href* attribute and the *DataContent* element are mutually exclusive; exactly one of the two must appear, but never both.

The News Content Format

All internally included news content items is in the Millistream News Content Format. The MIME type is “text/vnd.Millistream”, and the format is XML.

The Millistream News Content *DataContent* element contains up to five markup components:

- a required, non-repeatable *head* element containing header information
- an optional, non-repeatable *tagline* element containing the tagline for the news item
- an optional, non-repeatable *footer* element containing the news item footer
- an optional, non-repeatable *CompanyInfo* element containing summary information about the company sending out a press release
- a required, non-repeatable *body* element containing the main content of the news item in *xhtml* format

The *tagline* element contain plain text with no inline markup. The *footer* and *CompanyInfo* elements contains regular block content like paragraphs.

Header Information

The *head* element contains of up to four markup components:

- a required, non-repeatable *headline* element containing the headline and optional sub-headlines
- a required, non-repeatable *distributor* element containing the news agency
- an optional, non-repeatable *pubdata* element containing information about the item's publication
- an optional, non-repeatable *comment* element containing a comment from the publisher, this can be information on changes from a previous release or information about future followups.
- an optional, non-repeatable *abstract* element containing a summary of the news item

The *distributor* element contain plain text with no inline markup. The *headline* element contain exactly one *h1* element holding the main headline and zero or more *h2* elements holding sub-headlines: both of these contain regular inline markup (see below). The *abstract* element contain regular block content like paragraphs.

The *pubdata* element currently only contain the single attribute *ex-ref* containing the external reference to the rendering of the news object as published, in the form of a URL.

Block Content

Block content is information that is normally displayed starting on a new line, like paragraphs, lists, tables and so on. The News Content Format allows at least these elements to appear in a block context in any order, all optional and repeatable:

- the *p* element, representing a paragraph
- the *strong* element, representing strong emphasis (bold)
- the *em* element, representing weak emphasis (italic)
- the *u* element, representing underlying text
- the *h2* element, representing a sub-headline anywhere in the story
- the *ol* element, representing an ordered (numbered) list
- the *ul* element, representing an unordered (bulleted) list
- the *pre* element, representing preformatted text like an ASCII table or source code listing
- the *sup* element, representing superscript text

- the *table* element, representing tables

The *p*, *strong*, *em*, *u*, *ol*, *ul*, *pre*, *sup* and *table* elements are all compatible with HTML version 4 and XHTML.

The *p*, *strong*, *em*, *u* and *hl2* elements contain regular inline content, and the *pre* element contains plain text with no sub-elements.

Regular Inline Content

Most elements that allow text also allow inline elements to appear. We currently support five basic inline elements:

- *a*, for creating hyperlinks
- **img**, supplying links to images
- *q*, for tagging direct, inline quotations
- *virtloc*, for tagging URLs and e-mail addresses meant to be displayed
- *br*, for providing line breaks inside paragraphs

The *virtloc* element contains plain text (no inline markup) representing either an e-mail address or a World Wide Web uniform resource locator (URL):

```
<virtloc>testuser@millistream.com</virtloc>
```

This is an address intended to be shown rather than simply presented as a link. If the address contains the character “@”, it represents an e-mail address; otherwise it represents a URL.

The *a* element, from HTML, represents a World Wide Web hypertext link embedded in a document. The element has a required *href* attribute, containing the Uniform Resource Locator (URL) for the link:

```
<a href="http://millistream.com">Millistream AB</a>
```

The **img** element, from HTML, represents an external link to an image embedded in the document. The element has the required **src** and **alt** attributes containing the external link to the image and the alternative text for the image as well as the optional **height** and **width** attributes that provides the height and width of the image in pixels:

```

```

The *em* element, from HTML, marks an emphasized phrase. The element contains regular inline content:

```
<p>This is <em>really</em> what we wanted</p>
```

The *q* element contains an inline quotation without any kind of quotation marks added around it:

```
<q>This is all that we needed</q>, said Mr X, vice president of Y Corp
```

The *br* element, from HTML, represents a line break within a paragraph. Used when the lines should be displayed “closer together” than between paragraphs. The element will always appear closed:

```
<p>These rows<br/>Should all<br/>Appear on separate lines</p><p>This is a new paragraph</p>
```

This should be displayed something like this:

```
These rows
Should all
Appear on separate lines

This is a new paragraph
```

News Tags

News items can carry news tags in the **MDF_F_NEWSCODINGTAGS** field in MDF or the *tag* column in the *newstags* table in MDA. Supported values are (exactly which tags that are supported vary from source to source):

Content Type	
<i>tag</i>	<i>description</i>
exclusive	Interview or newsbeat
briefing	Market preview and summary
calendar	Corporate, economic and political event calendar
xtr	More reporting on the day's key company news
indepth	Deeper look at companies and trends
pickup	News reported by other media
memo	News Summary
commentary	Market Commentary
analytics	Company analysis
today	Today's company and/or market events
thisweek	Company and/or market events happening this week

Corporate Events	
<i>tag</i>	<i>description</i>
cos	All news about companies
report	Earnings report
yearend	Year-End report
annual	Annual report
interim	Interim report
estimates	Earnings Preview / Analysts Forecast
outlook	Earnings guidance by the company
sales	Sales and production updates
prodrel	New product announcements; product updates
tesres	Test results for products, services and resource deposits
exx	Executive changes
contracts	News about contracts and joint-venture agreements
regulation	News about company regulation and antitrust enforcement
litigation	Lawsuits and other legal actions or court disputes
regulatory	Regulatory news item
newsl	News Letter
gm	General Meeting
invitation	Invitation to shareholders meeting
research	Research report

nav	Reporting of Net Asset Value
-----	------------------------------

Corporate Actions	
<i>tag</i>	<i>description</i>
ca	All news about corporate actions
ma	Mergers & Acquisitions
issuance	Stock & Bond issuance, listing changes and other transactions
ipo	Initial Public Offer
dividend	Dividend declarations and other capital returns
insider	Share buying and selling by company insiders
financing	News about loans and other finance facilities
funding	Company funding, rights issues etc.
restructure	Company reorganizations

Market Activity	
<i>tag</i>	<i>description</i>
mkt	All news about market activity
stocks	Commentary on equities
kno	Deeper insight on focus stocks and special situations
smallcap	Insight on Small-cap stocks
rec	Buy-Sell-Hold recommendations; price target changes
mov	Share movers
idx	Changes to constituents of equity indices
bonds	Commentary on bonds
currencies	Commentary on currencies
cryptocurrency	Commentary on crypto currencies
commodities	Commentary on oil, gas and precious metals
mutualfunds	Commentary on mutual funds
hedgefunds	Commentary on hedge funds
etp	Commentary on exchange traded products
fixedincome	Commentary on fixed income

Macroeconomics	
<i>tag</i>	<i>description</i>
macro	All macroeconomic news
indicators	Economic indicators and reports
politics	Government decisions and policy
cbk	Central banks actions and policy
sov	Government borrowing

Industry	
<i>tag</i>	<i>description</i>
icb:<industry/super-sector>	Industry and Supersector from the Industry Classification Benchmark (ICB). E.g "icb:1000" for Basic Materials and "icb:1300" for Chemicals.
gics:<sector/group/industry/sub-industry>	Sector, Industry Group, Industry and Sub-Industry from the Global Industry Classification Standard (GICS). E.g "gics:10" for Energy, "gics:1510" for Materials, "gics:151010" for Chemicals and "gics:15101010" for Commodity Chemicals.
iptc:xxxxxxxx	The 8 digit IPTC Subject Code from http://cv.iptc.org/newscodes/subjectcode
agriculture	gics:20106015, gics:30202010
automotive	gics:2510, gics:25504050
aviation	gics:203010, gics:203020, gics:20305010
basemetals	
betting	
biometrics	
copper	
crowdfunding	
defence	gics:201010
ecommerce	
energy	gics:10
financials	gics:40
fintech	
foods	gics:302020
gaming	gics:25301010
gold	
healthcare	gics:35
industrials	gics:20
leisure	gics:252020, gics:253010
lifescience	gics:35202010
materials	gics:15
media	gics:2540
medtech	
oil	gics:10101010, gics:10102010, gics:10102020, gics:10102030, gics:10102040
oilservice	gics:10101020
pharmaceuticals	gics:35203010
precious	Precious Metals
realestate	gics:60
renewable	gics:55105020
retail	gics:2550

services	
shipping	gics:203030, gics:20305030
space	gics:201010
tech	gics:45
telecom	gics:50
transports	gics:2030
vr	Virtual Reality
wind	

Geography	
<i>tag</i>	<i>description</i>
geo:int	International
geo:ame	Americas
geo:apa	Asia Pacific
geo:eur	Europe
geo:eu	European Union
geo:bal	Baltic States
geo:mea	Middle East & Africa
geo:AF	Afghanistan
geo:AU	Australia
geo:BE	Belgium
geo:BR	Brazil
geo:CA	Canada
geo:CN	China
geo:CH	Switzerland
geo:DE	Germany
geo:DK	Denmark
geo:EE	Estonia
geo:ES	Spain
geo:FI	Finland
geo:FR	France
geo:GB	United Kingdom of Great Britain and Northern Ireland
geo:GR	Greece
geo:HK	Hong Kong
geo:IE	Ireland
geo:IN	India
geo:IS	Iceland
geo:IT	Italy
geo:JP	Japan

geo:LY	Libya
geo:NO	Norway
geo:NZ	New Zealand
geo:PK	Pakistan
geo:PL	Poland
geo:PT	Portugal
geo:RU	Russian Federation
geo:SE	Sweden
geo:US	United States of America
geo:ZA	South Africa

Custom tags

A news agency can send tags that are custom for that particular agency some news agencies might also send tags dynamically if a specific topic is of temporary interest.

Custom tags are prefixed with the symbol of the news agency (and a colon character) so as to not collide with the existing and defined news tags.

Examples:

FWN: *hedermeradagen*

ALL: *250*

RB: *Top/Web*

TT: *prod:TTEKO*

TT: *per:Håkan Nilsson*

TT: *org:Ikea*

TT: *place:Malmö, Sverige*

Examples

This section contains three real world examples of the Millistream News Format. Do please note that we have indented the examples for easy reading, in reality they are transmitted as compact XML (all on a single line with no space or tabs between the tags):

Example #1:

```
<?xml version="1.0" encoding="UTF-8"?>
<NewsItem>
  <NewsEnvelope>
    <TransmissionId>20090611:BIT:2059:0</TransmissionId>
    <DateAndTime>20090611T101846+0200</DateAndTime>
    <SentFrom>
      <Party FormalName="Stena Bulk"/>
    </SentFrom>
  </NewsEnvelope>
  <ContentItem>
    <MimeType FormalName="text/vnd.Millistream"/>
    <DataContent>
      <head>
        <headline>
          <h1>Stena Bulk öppnar kontor i Finland</h1>
        </headline>
        <distributor>Cision</distributor>
        <abstract>
          <p>Med de växande affärerna i den baltiska regionen och ytterligare
koncentration på mer isklassat tonnage är vi stolta att meddela att vi nu öppnar
```



Stena Bulk Finland med säte i Helsingfors.</p>

</abstract>

</head>

<body>

<p>Mauri Harki har utsetts till befракtningschef för det nya kontoret, som kommer att inrättas i centrala Helsingfors. Verksamhetens fokus kommer att vara trafiken i Östersjön och affärer med skandinaviska och ryska företag. Vårt nuvarande Moskvakontor flyttar sin verksamhet till Helsingforskontoret, som snabbt kommer att kunna täcka det ryska kontorets tidigare verksamhet. Stena Bulk London fortsätter att koncentrera sig på Nordsjö- och Medelhavsmarknaden.</p>

<p>Ulf G. Ryder, VD Stena Bulk säger: "Vi välkomnar detta nya steg som ytterligare stärker vår strategi att agera lokalt och nära kunderna. Stena Bulk Finland får genom sin placering en naturlig roll som omfattar de baltiska staterna och våra ryska affärer från ett strategiskt geografiskt helt rätt läge med folk som vet hur man bedriver verksamhet i dessa regioner."</p>

<p>Så snart det fulla teamet satts samman återkommer vi med namn/adresser mm.</p>

</body>

<footer>

<h2>Ulf G. Ryder</h2>

<p>VD
Stena Bulk AB
Tel +46 31 855001 direkt
Mobile +46 704 855001</p>

</footer>

<CompanyInfo>

<p>Stena Bulk - Med kontor i sju länder är Stena Bulk ett av världens ledande tankrederier. Genom bland annat en egen designavdelning levereras nyskapande lösningar för kundernas behov av transport och logistik. Företaget har en kontrollerad flotta på runt 75 tankfartyg och är aktivt inom alla delar av tankmarknaden. Stena Bulk har även marknadsförings- och driftsansvar för hela Concordia Maritimes flotta. Stena Bulk är en del av Stena Sfären som har drygt 17 500 anställda och en omsättning på 50 miljarder SEK. www.stenabulk.com</p>

</CompanyInfo>

</DataContent>

</ContentItem>

<ContentItem

Href="<http://feed.ne.cision.com/client/waymaker1/WOLReleaseFile.aspx?id=1229559&fn=release.html>" Role="External">

<MimeType FormalName="text/html"/>

</ContentItem>

<ContentItem

Href="<http://feed.ne.cision.com/client/waymaker1/WOLReleaseFile.aspx?id=1229559&fn=wkr0011.pdf>" Role="External">

<MimeType FormalName="application/pdf"/>

</ContentItem>

```
</NewsItem>
```

Example #2:

```
<?xml version="1.0" encoding="UTF-8"?>
<NewsItem>
  <NewsEnvelope>
    <TransmissionId>4281731</TransmissionId>
    <DateAndTime>20090611T092602+0200</DateAndTime>
    <SentFrom>
      <Party FormalName="Direkt-SE"/>
    </SentFrom>
  </NewsEnvelope>
  <ContentItem>
    <MimeType FormalName="text/vnd.Millistream"/>
    <DataContent>
      <head>
        <headline>
          <h1>*** HEMTEX: HAR INTE SETT ANLEDNING ÄNDRA LÅNGSIKTIGA FIN-MÅL-FC</h1>
        </headline>
        <distributor>Direkt-SE</distributor>
      </head>
      <body>
        <pre>HEMTEX: HAR INTE SETT ANLEDNING ÄNDRA LÅNGSIKTIGA FIN-MÅL-FC</pre>
      </body>
    </DataContent>
  </ContentItem>
</NewsItem>
```

Example #3:

```
<?xml version="1.0" encoding="UTF-8"?>
<NewsItem>
  <NewsEnvelope>
    <TransmissionId>1322061</TransmissionId>
    <DateAndTime>20090611T100149+0000</DateAndTime>
    <SentFrom>
      <Party FormalName="Downing Protected VCT I PLC"/>
    </SentFrom>
```



```
</NewsEnvelope>
<ContentItem>
  <MimeType FormalName="text/vnd.Millistream"/>
  <DataContent>
    <head>
      <headline>
        <h1>Transaction in Own Shares</h1>
      </headline>
      <distributor>Hugin</distributor>
    </head>
    <body>
      <pre>Downing Protected VCT I plc
Transaction in own shares
11 June 2009
```

Downing Protected VCT I plc announces that on 11 June 2009, the Company purchased 20,000 ordinary shares of 50p each for cancellation representing 0.25% of the issued share capital at a price of 82.0p per share.

---END OF MESSAGE---

This announcement was originally distributed by Hugin. The issuer is solely responsible for the content of this announcement.</pre>

```
</body>
  <tagline>Copyright © Hugin AS 2009. All rights reserved.</tagline>
</DataContent>
</ContentItem>
</NewsItem>
```

Changelog

- 2018-11-22** *Added info about the News Tags*
- 2015-10-21** *Added the **u** element to the Block content.
Added the **img** element to the Regular Inline content.*
- 2015-02-19** *Added the **Role** attribute to the **ContentItem** element.*
- 2012-09-12** *Added the **comment** element to the **head** element.*
- 2011-09-09** *Added the **strong** element to the Block content.
Moved the **em** element from Regular Inline content to Block content.*
- 2011-01-26** *Added the **pubdata** element.*
- 2009-06-11** *Initial version.*