ABC ERRORS

errors (critical)A

These errors severely impact the possibility for customers to find information or receive a hit in their stored search profiles.

MA A large part of the *Missing Article*

article text is missing (in its own article or as a part of a combined article) according to the following criteria:

> A) More than 100 words of the article are missing.
>
> or
>
> B) More than 30% of words in the article are missing, but minimum 5 words.
>   If less than 5 words are missing, it's a type B error. (This situation concerns very short articles, often found on the front page).

> The article word count is based on words in all main text elements, i.e., headline, subheadline, pre-intro, intro, body, caption, quote, blurp, and fact box.

B errors (major/serious)

These errors still have a big impact on customers' ability to find information or receive a hit in their stored search profiles.

MT *Missing text*
Text is missing from the headline, subheadline, pre-intro, intro, body, quote, blurp, or fact box, but not enough text is missing for the error to be a type A error.

Text missing in more than one article text element still counts as one error. E.g., if the text is missing in the headline and in the story, then it counts as one B-error.
Note that missing text from an image caption is NOT an MT error, but a WIC error (see below).

IPA *Incorrectly Placed Article*
The reference to the physical position of the page is wrong (e.g., we indicate that the article is on pages 4 and 5, but in reality, it is on pages 8 and 9).

SICW *Split or Incorrectly Concatenated word*
> A) Extra space characters exist that cause one or more words to split
>
> Or

B) One or more missing space characters cause words to merge.

Example 1: "Flo wer" instead of "Flower".
Example 2: "Youhave a flower" instead of "You have a flower".

HW *Hyphenated Words*
Wrong use of hyphenation. Keeping a hyphen when it should be left out or removing it when it should be kept in.

Example 1: "Flo-wer" instead of "Flower".
Example 2: "66årig" instead of "66-årig".

WB *Wrong Byline*
The byline tag is provided wrongly.

A) Byline name, email, phone, Twitter handle, or accompanying info (e.g. journalist title) is missing, misspelled, or put in the wrong (byline) element.
B) Non-byline content is present in the byline.

IDA *Incorrectly Divided Article*
The article is split into two or more separate articles.

JC *Jumbled Content*
A) Jumbled words, sentences, or paragraph

All words are there and correctly spelled but they are placed in the wrong order in the XML / JSON element.

Smaller or larger parts of the text are placed in the wrong order relative to the article, causing problems for the end customer to understand the article and/or for Media Track's client to search based on the distance between keywords.

Example 1 (wrong word order): "flower have a You " instead of "You have a flower".

Example 2 (wrong sentence order): "I have a flower too. You have a flower." should have been "You have a flower. I have a flower too."

B) Jumbled images, graphs, or visual elements

Visual elements are not in the right position in the cutout relative to their position in the

original article. This is mostly relevant for rearranged and reflowed
cutouts.

WTC   *Wrong Text Character*
One or more misspelled or missing letter(s) in word(s).

Includes combination letters (e.g. letters that newspapers concatenate into a single character.
Examples are "fj", "ff" etc.).
Note that, e.g., WTC is not used for characters, which are not inside words. E.g., a "#" which
should have been a "*" is NOT a WTC error.

Example 1: "Ratt" instead of "Rätt".
Example 2: "Efect" instead of "Effect"

WIT   *Wrong Illustration Text*
This includes problems with captions and graphical element text

A) One or more image caption text(s) are missing, misspelled, or refer to the wrong image.
B) Missing text from graphical elements like graphs or maps with text. Includes
   situations where tables have a graphical element in the background.

WGE   *Wrong Graphical Element*
One or more graphical elements, i.e., images or illustrations, are wrong or missing.

Includes wrong image, cut image or illustration, and missing background image on tables or
factboxes

WCL   Wrong Cutout Layout

C) ~~One or more image caption text(s) are missing, misspelled, or refer to the wrong
   image.~~
D) ~~Missing text from graphical elements like graphs or maps with text. Includes
   situations where tables have a graphical element in the background.~~

A)   Layout has issues to the extent that the article cannot easily be read
For example, tiny parts of characters are missing, text is too small to be read, large variations in
column positions.

B)   Layout is not aesthetically pleasing
For example, Additional spaces, tiny remains of characters / illustrations from surrounding

articles, uneven font sizes

C errors (issue)

These errors will have less impact on customers' ability to find information or receive a hit in their search profiles. The errors include issues that may impact the presentation of the article to the end customer.

IIA *Incorrectly Included Article or Text*

Unwanted information is indexed as article information. Includes the case where text, which is not supposed to have been plotted, is added to an article.

Such information/article is typically an ad or a text which guides the reader in how to send a letter to the editor.

CA *Concatenated articles*

Articles combined into one article.

WT *Wrong Tag*

Text exists, but in wrong XML element/attribute, or repeated in multiple XML elements. Exception for byline name, byline name info (e.g., journalist title) and image text which are B type errors.

Example 1: The article headline "White House Scandal" was wrongly put as subhead and the subhead "Clinton resigns as president due to new rumours" was wrongly put as headline.

WLB *Wrong Line Break*

Line break errors.

WTF *Wrong Text Format*

Upper and lower case errors in text.

Example. "tAnyA mAdrId" instead of "Tanya Madrid"

WNTC *Wrong Non-Text Character*

Wrong or missing dots, bullet points, squares and symbols in lists and text areas.

Example. A "#" which should have been a "*" is a WNTC error.

MPC   *Missing Punctuation Character*

Punctuation characters include, e.g., comma (,), semi colon (;), colon (:), full stop (.), exclamation mark (!), and question mark (?).

Example. "Give me that" should have been "Give me that!". I.e., exclamation mark missing.

PR   *Problems with Photographer/ Photographer rights*
Image photographer or rights are missing, misspelled or corrupted.

MOAS   *Missing Or Additional Space*
Missing space characters or extra space characters where error does not lead to splitting or merging of words. Threshold is set to three or more trailing spaces to be counted as a C error.

Example 1: "You knew   that!" instead of "You knew that!"
Example 2: "-We left early" instead of "- We left early". I.e., missing space after hyphen.

WSN   *Wrong Section Name*
Logical_position has wrong reference. I.e., section name is either missing or incorrect.

Example: Section name has wrong letter casing (fx "Nyheter" when supposed to be "NYHETER").

Out-of-category issues (not an error)
- Subheadline and pre-intro is used as support elements when text is not a headline or intro, but something in between. We can tolerate some variances in the use of these elements, but when both are used we request that the most headline-like text is put in the subheadline element.
- Dots, squares and symbols in beginning of headline and byline name can be omitted.