

NULL値定数変換のバージョン差異

2025-11-20

Pentaho 9とPentaho 10の仕様差異を日々見たのでメモ

概要

デフォルト設定(オプション `KETTLE_EMPTY_STRING_DIFFERS_FROM_NULL=N`) の状態で、 Pentaho 9 と Pentaho 10 の仕様差異がステップ **NULL値定数変換** で見られた。

この定義で対象項目(この例ではワード2)がnullだった場合の振る舞いに差異があった。

「置き換える値」には何も定義しない。

ステップ名

すべてのフィールドのNullを置き換える。

置き換える値

空文字を設定

マスク(日付)

フィールドを選択

値のタイプを選択

値のタイプ

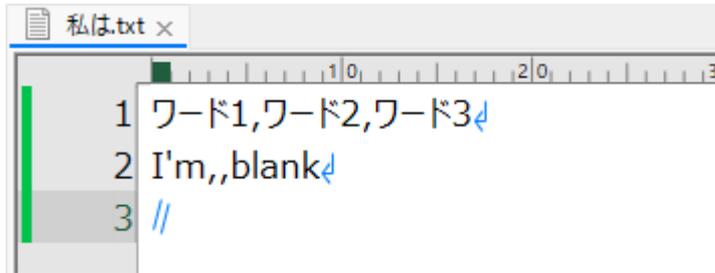
#	タイプ	置き換える値	変換マスク(日付)	空文字を設定

フィールド

#	フィールド	置き換える値	変換マスク(日付)	空文字を設定
1	ワード2			N

ヘルプ OK(O) フィールドを取得(G) キャンセル(C)

入力データはこれ。

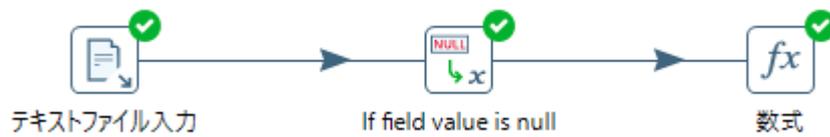


変換結果はこうなる。

バージョン	変換結果
Pentaho 9	null
Pentaho 10	長さゼロ文字列

定義

データ変換定義



文字列として読み込みする。

テキストファイル入力													
ステップ名 テキストファイル入力													
#	フィールド名	データタイプ	書式	位置	長さ	精度	通貨記号	桁区切り文字	数値囲み文字	NULL可能	デフォルト	空白除去	データを代替する
1	ワード1	String						.	,	-		なし	N
2	ワード2	String						.	,	-		なし	N
3	ワード3	String						.	,	-		なし	N

項目ワード2の定義はするけど置換先のデータは登録していない。

NULL値定数変換

ステップ名

すべてのフィールドのNullを置き換える。

置き換える値

空文字を設定

マスク(日付)

フィールドを選択

値のタイプを選択

値のタイプ

#	タイプ	置き換える値	変換マスク(日付)	空文字を設定

フィールド

#	フィールド	置き換える値	変換マスク(日付)	空文字を設定
1	ワード2			N

項目ワード2がブランク判定なら“so”非ブランク判定なら“not”に置き換えする。

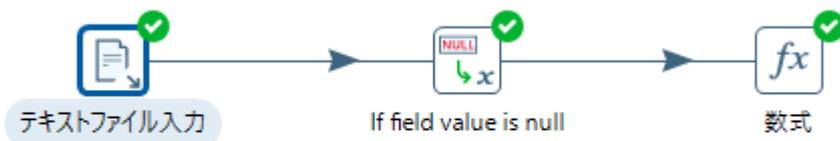


実行結果(Pentaho 9)

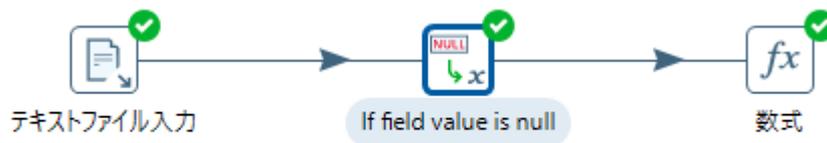
デフォルト状態

デフォルト状態(オプション KETTLE_EMPTY_STRING_DIFFERS_FROM_NULL=N)で実行。

項目ワード2はnullとして取り込まれている。



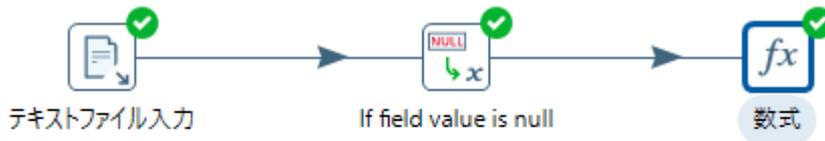
変換先の定義が無かったのでnullだった場合はそのままnullになっている。



実行結果

#	ワード1	ワード2	ワード3
1	I'm	<null>	blank

数式内のISBLANK([ワード2])の結果がTRUEになる事がわかる。つまりnullはblankと判定される。



実行結果

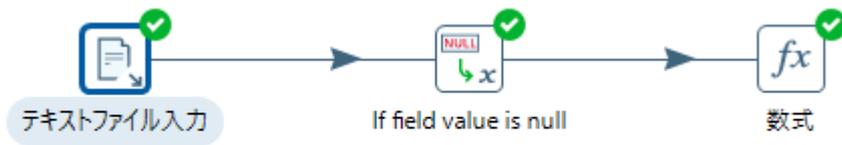
#	ワード1	ワード2	ワード3
1	I'm	so	blank

KETTLE_EMPTY_STRING_DIFFERS_FROM_NULL□Y

オプション KETTLE_EMPTY_STRING_DIFFERS_FROM_NULL=Y を設定して実行してみる。

#	変数名	値	説明
28	KETTLE_DEFAULT_SERVLET_ENCODING		Defines the default encoding
29	KETTLE_DEFAULT_TIMESTAMP_FORMAT		The name of the default timestamp format
30	KETTLE_DISABLE_CONSOLE_LOGGING	N	Set this variable to prevent Kettle from logging to the console
31	KETTLE_DO_NOT_NORMALIZE_NULL_STRING_TO_EMPTY	N	Prevents Kettle from normalizing null strings to empty
32	KETTLE_DO_NOT_NORMALIZE_SPACES_ONLY_STRING_TO_EMPTY	N	Prevents Kettle from normalizing spaces-only strings to empty
33	KETTLE_EMPTY_STRING_DIFFERS_FROM_NULL	Y	NULL vs Empty string comparison
34	KETTLE_FAIL_ON_LOGGING_ERROR	N	Set this variable to fail on logging errors

項目ワード2がnullではなく（おそらく）長さゼロの文字列として読み込まれている。



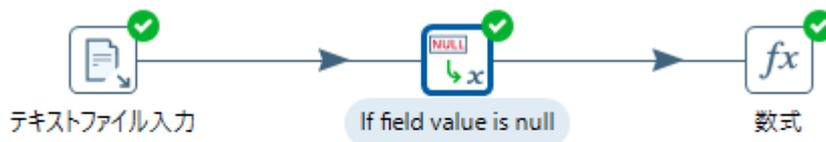
実行結果

ログ | 実行履歴 | 実行状況 | パフォーマンスグラフ | メトリクス | データをプレビュー

開始行 最終行 Off

#	ワード1	ワード2	ワード3
1	I'm		blank

項目ワード2はnullではないから変換もされない。



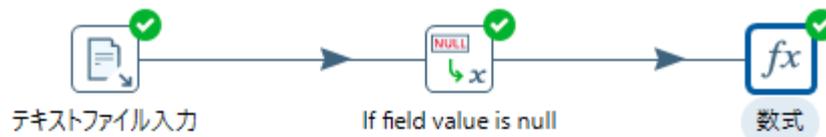
実行結果

ログ | 実行履歴 | 実行状況 | パフォーマンスグラフ | メトリクス | データをプレビュー

開始行 最終行 Off

#	ワード1	ワード2	ワード3
1	I'm		blank

数式内のISBLANK([ワード2])の結果がFALSEになる事がわかる。つまり長さゼロの文字列は非ブランクと判定される。



実行結果

ログ | 実行履歴 | 実行状況 | パフォーマンスグラフ | メトリクス | データをプレビュー

開始行 最終行 Off

#	ワード1	ワード2	ワード3
1	I'm	not	blank

実行結果(pentaho 10)

今現在、個人で Pentaho 10 を取得するのはハードルが高いので、実行結果等の例示できない。ごめん。

最近仕事で、類似の定義を扱った。

オプションがデフォルト状態でデータ変換定義を実行したところ、項目ワード2 が null の時 NULL値定数変換 を通すと長さゼロの文字列に無条件に変換される事象を経験した。

意図としては何か要件が増えた時用に配置しておいて通常は無変換にしたかったんだと思われる。

Pentaho 9からのPentaho 10への移行で、無変換を期待したのに、無条件に変換される事になってしまった。

結構な影響がある状態。青くなった。

[技術資料](#), [etl](#), [pentaho9](#), [pentaho10](#), [NULL値定数変換](#), [仕様差異](#)

From:

<https://wiki.hgotoh.jp/> - 努力したWiki

Permanent link:

<https://wiki.hgotoh.jp/documents/etl/pentaho/pentaho-013>

Last update: **2025/11/20 00:22**

