

「DAILY YOMIURI 記事データ集」tag/データ構成

1. ファイルについて

- ・文字コードは SJIS です。
- ・DOS 用改行コード(CR,LF)ではなく、UNIX 用の改行コード(LF のみ)を使用しております。

2. データ形式

- ・形式は “ tag:データ ”
- ・tag/データともにすべて半角です。
- ・ファイルは 1 年で 1 ファイルとなります。

3. tag/データ構成

データ項目	TAG	データ内容/備考
ID 番号	ID	数字 8 桁 ユニーク
記事 ID	C0	例) 20030101TDY01001 ユニーク
掲載年月日	YF	例) 20030101
頁	Y0	例) 01
面名	Y7	データ内容は下記参照
タイトル	T1	記事見出し
記事本文	T2	記事本文

4. データ内容

1) 記事 ID(tag.C0)

- ・データ内容は以下の通りです。
掲載年月日(8 桁) + 紙誌名(3 桁) + 頁(2 桁) + ページ内記事通番(3 桁)

2) 紙誌名(tag.C0 内)

- ・TDY は THE DAILY YOMIURI の略です。(東京本社朝刊のみの発行となります)

3) 面名(tag.Y7)

- ・データ内容(種類)は以下の通りです。
2002 年以前はデータが入っていない場合があります。

Advisories	Financial,econo	Science
Books	Food	Society
Comentary	Gnrl,Dfens,Ecno	Special Events
Crime-Accident	Imperial Family	Sports
Culture	Industry	Sports score
Domestic	International	Standing
Economy	Lifestyle	State,regional
Entertainment	Political	Travel
Financial pages	Politics	Washington
Financial table	Racing results	Weather