surfaceForm: hiển thị

reading: cách đọc

baseform: nguyên mẫu

partofspeech: từ loại

position: vị trí (id)

２９日の朝, **インドネシア**の**ジャカルタ**を出発した「**ライオン・エア**」の飛行機が海

に落ちました。この飛行機は**インドネシア**の[観光地](javascript:void(0))の**バンカ島**に向かっていました。

消防によりますと、31日午前10時半ころ愛知県豊明市の豊明学園豊明幼稚園で、「園内で異臭がして、複数の人がめまいと吐き気を訴えている」と通報がありました。

消防が駆けつけたところ、園児らがせきや頭痛、おう吐などの症状を訴えていて、子ども９人を含む、合わせて23人が病院に搬送されたということです。  
消防によりますと、いずれも意識はあるということです。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Kuromoji** | **Mecab** | **KyTea** | **Juman** |
| **Word segmentation** | **X** | **X** | **X** | **X** |
| **Part-of-speech tagging** | **X** | **X** | **X** | **X** |
| **Lemmatization** | **X** | **X** |  |  |
| **Readings** | **X** | **X** | **X** |  |

Chức năng:

* In flashcard
* Tăng tính năng (hình ảnh, âm thanh,...)

Lưu ý:

* Giải thích cho người nghe về 2 bảng chữ cái, về kanji
* Đối tượng người sử dụng
* So sánh tốc độ, độ chính xác giữa các tool (đọc bài báo tiếng Nhật)

Xem lại bố cục slide thuyết trình, hướng dẫn ghi tài liệu tham khảo

* Word segmentation - segmenting text into words (morphemes)
* Part-of-speech tagging - assign word-categories (nouns, verbs, particles, adjectives, etc.)
* Lemmatization - get dictionary forms for inflected verbs and adjectives
* Readings - extract readings for kanji