

Input free text
(e.g. GP history
or physical exam,
scanned clinic
letters, discharge
summaries)

Identify dates, numbers
and words which may
be part of clinical terms.
Attempt to correct
spelling mistakes

Search for word patterns
to assign attributes

Extend attributes (e.g. negation)
to a sequence of words, guided by
punctuation and stop words (e.g. 'but')

Attempt to match
sequences of words
to Read terms

Output structured data

Example analysis

Fracture of right femur in Jul 2010

fracture (WORD)
of (IGNORE)
right (WORD)
femur (WORD)
in (IGNORE)
jul 2010 (DATE: July 2010)

Pattern matching:
'... in (DATE)' maps to '... in jul 2010'
means that the date refers to the
preceding diagnosis

Matching to Read/OXMIS term

'fracture of right femur' maps to 'fracture femur'
(OXMIS term 8210 'FRACTURE FEMUR')

'of' is ignored

'right' is ignored as long as there is no 'left' in the Read term

1. OXMIS 8210 FRACTURE FEMUR
2. Date of preceding diagnosis: July 2010