Wikipedia人的資源マイニング

プロジェクトマネジメントコース　矢吹研究室　1142066　曽我勇貴

1．研究背景

　Wikipediaはウェブで最も有名なオンライン百科事典であり，今日では，執筆者の協力を失いつつある現状である[1]．このオンライン百科事典の英語版は2007年から2012年までに執筆者の30％を失っている．(現在3万1000人．世界のWikipediaの総計は25万人[1]．)

　Wikipediaは非営利組織であり，2001年1月15日に創設され，2003年6月20日に「ウィキメディア財団」がジミー・ウェールズにより創設された．

　財団の目的は，ウィキを用いて，不特定多数のボランティアによりオープンコンテキストの知的資源を開発することである．知的資源は無料で一般に提供されている．

　Wikipediaでは，登録者数がほとんど変わっていない[2]ので，Wikipediaの日本語版の統計[3]によると，日本の4年間の統計では，「活動中の登録者数」は，2009年11月1日の11,310人に対して，2013年12月12日は，10880人と減少している．また，Wikipediaには，洗練された法人が定められていないために，新しいことをしないのが現状である．

　Wikipediaは誰でも投稿・編集できるウェブ百科事典であるため，内容の信頼性を疑問視する声もある．また，記事を誰でも投稿ができるために，問題のある記述がなされた場合，投稿・編集された記事を善意の人に一任される．完全な自由主義なため悪意のある書き手を防ぎきれないという指摘がある．記事は完成・確定されることはないため，新しい情報にいつでも改変する事ができる．

2．研究目的

　Wikipediaを一つのプロジェクトとみなし，このオンライン百科事典で人的資源がどのように活用されているかを調査する．この調査により，オープンな共同作業プロジェクトにおける，人的資源マネジメントのあり方についての知見を得たい．

3．研究方法

　Wikipediaの作業履歴データを取得し，どのように人的資源が活用されているかを明らかにする．また，オープンな共同作業プロジェクトにおける人的資源マネジメントのあり方を提案する．

4．成果物のイメージ

　閲覧数の多い記事の編集者数と総編集回数のヒストグラムを書き，傾向に偏りがあるかを調査し分類する．総論．

5．進捗情報

　Wikipediaの登録者数の変動について，Wikipediaの参加者の数値が急激に変化していたところをGoogle Trendを使い調査したところ，同じような傾向が見られ，SOPA法案への抗議のためのものによると考察した．

　Wikipediaのグラフ[4]を読み取り，Excelのピポットテーブルを用いてグラフと表を作成した．

Wikipediaマイニング[5]を調査し，WikipediaをWebマイニングしたものであり，Webコーパスとして有益な情報を抽出する手法がある．総論．

6．今後の計画

　Wikipediaの全データをダウンロードしたものを，APIで解析する．

　Wikipediaのコミット回数のひすとぐらむを描き，どのような傾向がみられるか調査する．

参考文献

[1]“Wikipediaが岐路に立っている？”

<http://wired.jp/2013/11/27/the-decline-of-wikipedia/>(参照2013.11.2).

[2]“ウィキペディアの衰退”

<http://www.technologyreview.com/featuredstory/520446/the-decline-of-wikipedia/>(参照2013.10.23).

[3]“Wikipedia：日本語版の統計”

<http://ja.wikipedia.org/wiki/Wikipedia:%E6%97%A5%E6%9C%AC%E8%AA%9E%E7%89%88%E3%81%AE%E7%B5%B1%E8%A8%88>(参照2013.11.4).

[4]“ウィキペディア　多言語統計　日本語版”

<http://stats.wikimedia.org/JA/ChartsWikipediaJA.htm>

(参照2014.9.3).

[5]“Wikipediaマイニング”

<http://sigwp.org/ja/index.php/Wikipedia%E3%83%9E%E3%82%A4%E3%83%8B%E3%83%B3%E3%82%B0>(参照2008.2.18).