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，多くの人がボランティアで執筆するオンライン百科事典プロジェクトである．

Wikipediaは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．現在2014年9月1日では，287言語のWikipediaが開設されており，この全てのオープンコンテキストの知識資源は無料で一般に提供されている．

多くの人が参加するプロジェクトの代表例であるWikipediaを調査することによって，このような形式のプロジェクトのマネジメントについての有意義な知見が得られることが期待される．

ウィキペディアの成功理由についてはさまざまな考察がされている．例えば，「Google」の検索結果の上位(最初のページ)に「Wikipedia」が表示されていることが多い．これは，ウェブ百科事典として，検索に使う人が多くなり，新しい情報が入手しやすくなったと考察する．

しかし，一部の記事によるとWikipediaが衰退してきていると書かれている．Wikipediaのオンライン百科事典の英語版では，執筆者の協力を失いつつあり，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の日本語版の統計[2]を見ても同じような傾向がみられ，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．

本研究では，参加者の振る舞いのデータをマイニングすることによって，Wikipediaの成功理由を見つけることを目指す．また，衰退してきているWikipediaの執筆者の離脱を防ぐ手段をとることができるのかを考案する．

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[3]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[4]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のコミット回数のヒストグラムを描き，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]“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).

[4]“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).