

# Web サービスの障害がプロジェクトに与える影響を Twitter を活用して調査する方法

プロジェクトマネジメントコース 矢吹研究室 1442012 岩瀬翔

## 1. 背景

複数のメンバが同時に開発を行うソフトウェア開発プロジェクトでは、Web サービスを使うことがある。例えば、チーム内でファイルのバージョンを管理することのできる「GitHub」というサービスがある [1]。その GitHub のサーバーが 2016 年 1 月 28 日にダウンし、インターネット上で話題になった [2]。

GitHub のサーバーがダウンしてしまった実例から、プロジェクトにも影響を及ぼすのではないかと考えた。したがって、GitHub で障害が発生した場合どのような影響が発生しているのか調査する。

## 2. 目的

本研究の目的はソフトウェア開発プロジェクトで使用される Web サービスに障害が発生した場合、どのような影響が発生し、その影響がどれほどの人に及ぶのか調査すること。また、調査結果から障害発生に対するリスク対策案を検討することである。

## 3. 手法

調査方法は Twitter で投稿されている障害発生に関するツイートをデータとして収集する。Twitter では API が提供されているが、仕様によって過去のデータが取りきれないので以下の手順で行う [3]。

1. Twitter の検索結果を自動的に一番下までスクロールして、そのページ全体を HTML で保存するプログラムと HTML ファイルからツイートの時間と本文のみを抽出するためのプログラムを作成する。
2. GitHub に関連するすべてのサービスを継続的に状況監視している「GitHub Status」を参照し、2016 年で主要なサービスが停止、復旧したと記録されている時間を調べる。その時間前後を作成したプログラムを用いて Twitter で検索し、ページ全体を HTML で保存・データ抽出をする。これを各障害の発生時間ごとに繰り返す。

3. 障害発生から復旧までの GitHub に対するツイート数と、どのくらいの時間で復旧が完了するのかを調べる。調査結果をもとに、リスク対策案を検討する。

## 4. 想定される成果物

GitHub における障害発生から復旧までに投稿されたツイート数のグラフ及び、障害の発生から復旧までの時間でグラフを作成する。これらのグラフから考察することにより、障害発生に対するリスク対策案を検討する。

## 5. 進捗状況

GitHub Status を参照に調べたところ、2016 年の GitHub における障害発生回数は 14 回であった（ただし、10 月 21 日の障害は同時に Twitter もダウンしていたため対象に加えない）。それらの障害についてはデータの収集を完了し、障害の発生から復旧までに投稿されたツイート数のグラフ及び障害の発生から復旧までの時間でグラフを作成することができた。

## 6. 今後の計画

今回取得したツイートが、Twitter にて販売されているデータと比べ、抜けや間違いがないか調査することを検討している。

## 参考文献

- [1] 千葉祥太郎. 開発現場でホントに役立つツール (第 1 回) バージョン管理ツール「git」ソースコードは個人で管理 柔軟にワークフローを構築. 日経 systems, No. 252, pp. 92–97, apr 2014.
- [2] 岩城 俊介. Github がダウン、「ぜんぶのせいだ」のような惨状 今後の課題も. <http://breaking-news.jp/2016/01/28/022573> (2016.06.29 閲覧)。
- [3] 鳥海不二夫. Twitter 上のビッグデータ収集と分析. 組織科学, Vol. 48, No. 4, pp. 47–59, 2015.