

ソフトウェア開発の議論における 情報統合の提案

学籍番号：1421030 氏名：石川 俊明

指導教員：鷹野孝典 准教授

研究背景

- ソフトウェア開発では、設計や実装、技術に関する情報共有・意見交換をする必要が必ずあり、複数の手段を利用して頻繁におこなわれている。
- 会話や議論の情報は個人でのメモや開発用ツールに蓄積されているものの、同じ組織内で同様の話を繰り返す事がある。



関連研究

1. 『知識活動支援システムによる会議コンテンツ間の関連性の獲得とその応用』
(2010,情報処理学会 全国大会講演論文集)
2. 『議論マイニングによる議論掲示板利用者の能力推定』
(2017,情報処理学会 全国大会講演論文集)
3. 『参加者の議論能力と役割を考慮したオンライン議論の分析』
(2013,人工知能学会 全国大会論文集)
4. 『係り受け構造を用いたWeb議論掲示板における投稿への自動分類』
(2016,電子情報通信学会 技術研究報告)

研究課題

- 様々な手段で行われた会話・議論で生まれた情報をどのような手法を用いて統合するか.
- 発言・議論内容の関連度をどのように抽出するか.

研究動機

- 実際にソフトウェア開発に関わった時に「なぜこの技術なんだっけ？」「この人に実装を任せれば早く終わったのに」といった振り返りをする事があった.
- もし、様々な場所での会話が統合されていれば、こういった問題を解決出来ていたのではないかと考えた.

本研究の目的とアプローチ

- **目的**：議論手段に依存しない議論情報を集約可能な手段の検討と開発
- **アプローチ**：議論がどこで行われたかを指定し、発言内容にタグを付けて記録する。
 - 発言記録時に、その発言のトピック（カテゴリ）を抽出し、そのカテゴリに関連する発言・議論を推薦する。
 - 過去の議論の情報を提供することで、情報を持つユーザーの推薦や、議論の元となった問題の解決を図るシステムを提案する。

議論をタグに置き換える例

- ゼミの場合

- 「研究のアイデア」
- 「実装のアイデア」
- 「情報共有のあり方」

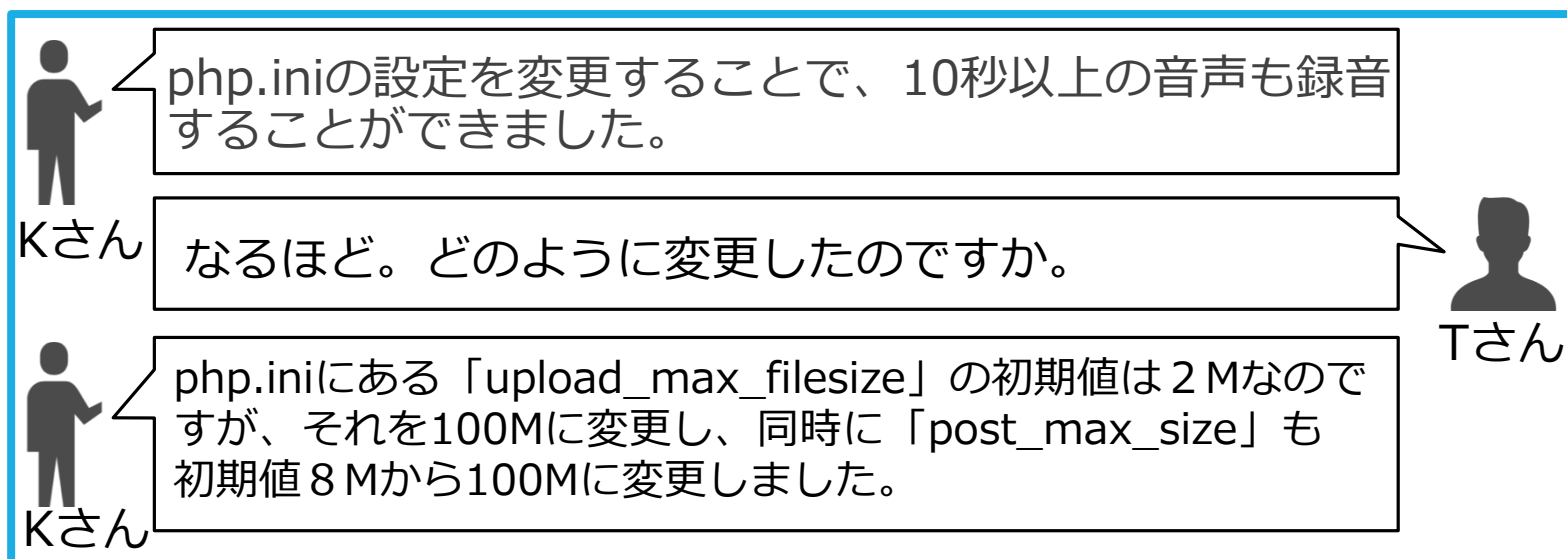
→タグとして考えると「研究」「実装」「アイデア」「アドバイス」などが挙げられる。

- 勉強会の場合

- 「PHPの使い方」
- 「JavaScriptの使い方」
- 「Slackの使い方」
- 「GitHubの使い方」

→タグとして考えると「プログラミング」「Webサービス」が挙げられる

実際の議論（会話）例



- 上記例の場合、「**PHP**」や「**サーバー**」，「**解決**」といったタグで記録する。これで，他に議論をした時に情報として推薦する事が出来る。
- 「**サーバー**」に関して知識のある人物の発見も可能と考えられる。

従来方式との比較

【メモで記録する場合の特徴】

- アイデアそのものや決定した方式そのものが記録される.

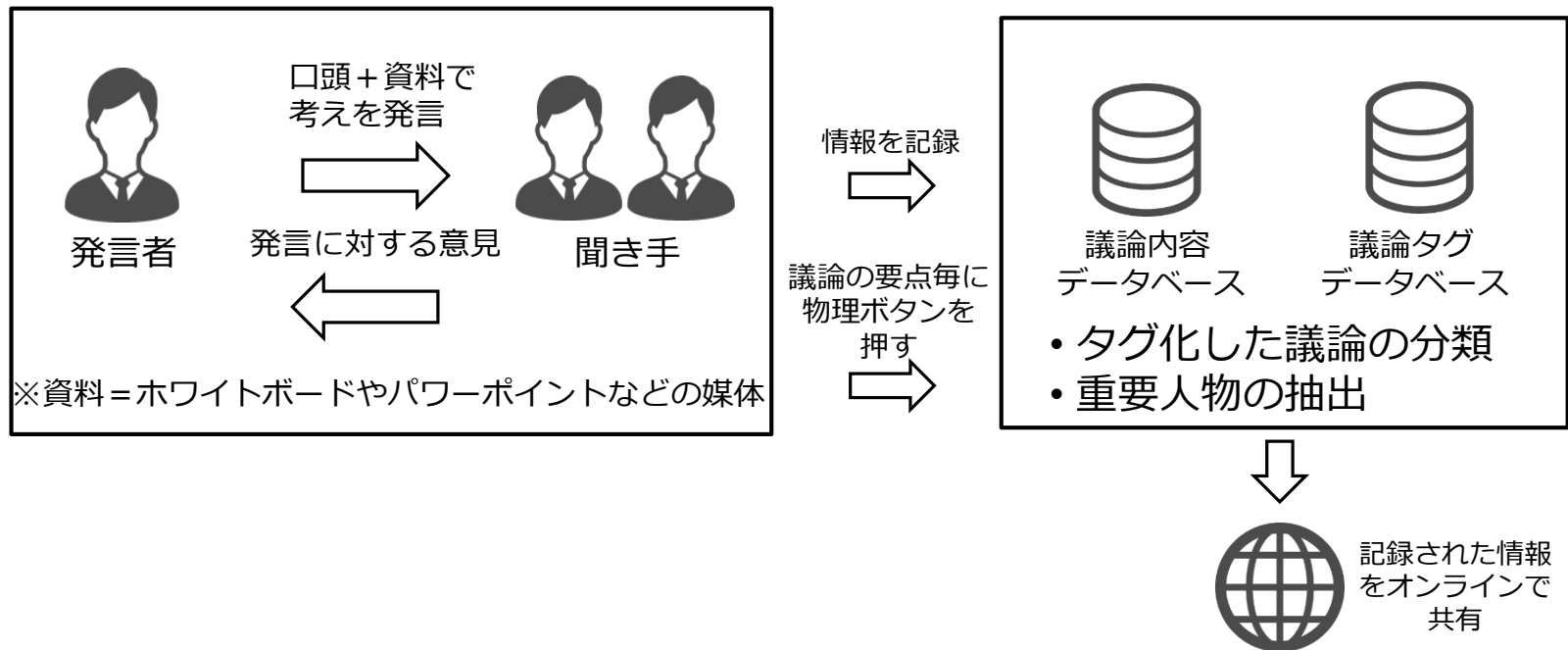
【本提案方式の特徴】

- アイデアを提案した人や実装方式に関する一連の議論が記録される.

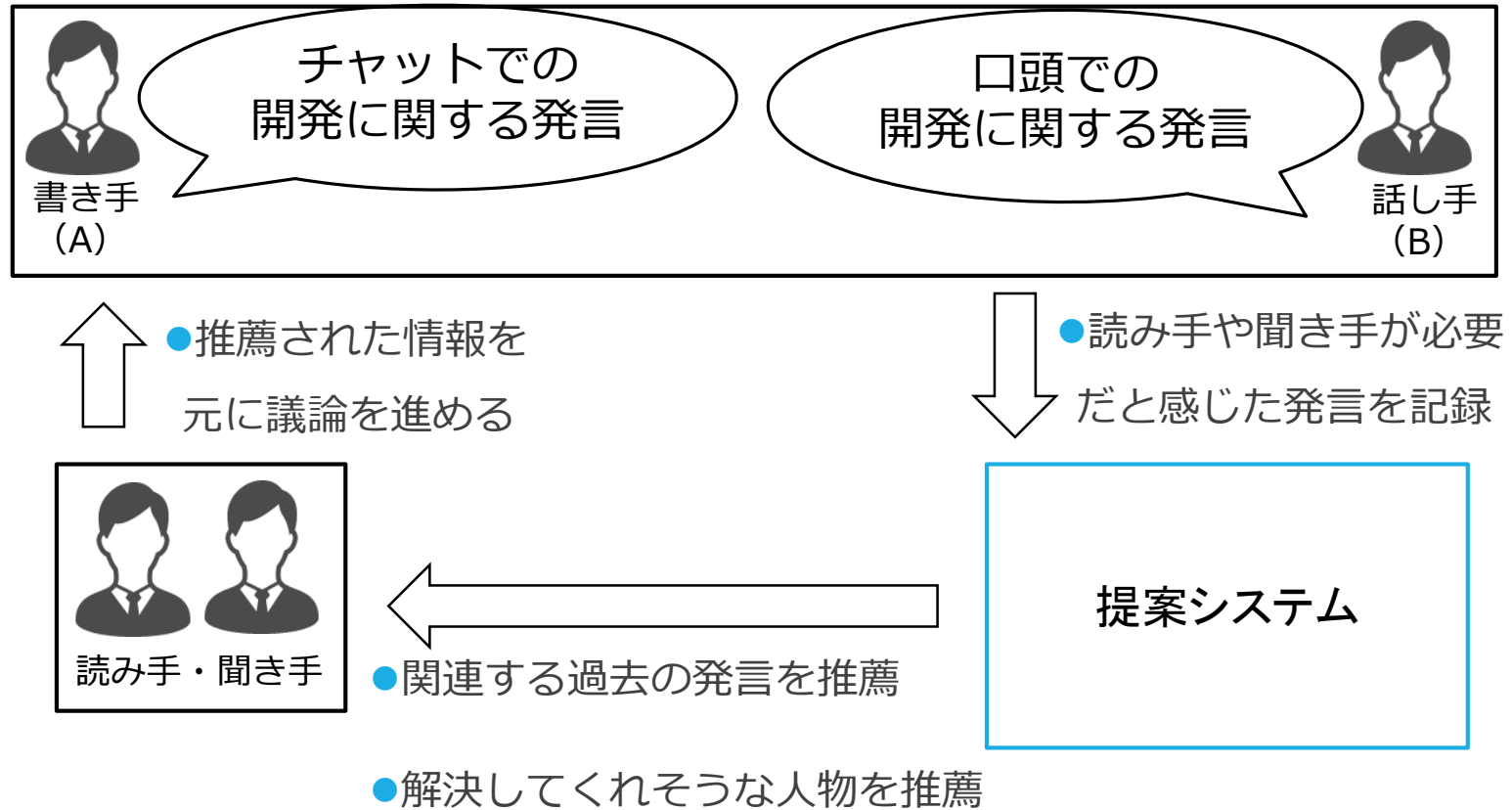
→ 本提案方式によって「アイデアを提案した人物」や「実装方式の利点」が記録される.

提案システム図

議論の場



提案システムの流れ[WIP]



実装

- Slackの1発言に対してタグを付けて記録するためのデータベースとプロトタイプシステムの作成

tags table

No	column name	内容
1	id	タグID
2	name	タグ名

users table

No	column name	内容
1	id	ユーザーID
2	name	ユーザー名
3	password_digest	パスワード

documents table

No	column name	内容
1	id	発言ID
2	author_name	発言者名
3	desc	発言内容
4	source	発言元
5	said_time	発言時間

- 中間テーブルを作成し、誰がデータを追加したか、発言にどんなタグが付けられたかを紐付ける

実装 (2)

ようこそ ishikawaさん
記録ページ
チャンネル名を指定する

2017-sotsuken ▼

タグを指定する

☐ Ruby ☐ C++ ☐ Python ☐ 未解決 ☒ 解決 ☐ Server

記録用ページ

確認

チャンネル名 2017-sotsuken
つけるタグは「解決」です。

発言者	発言内容	発言時間
tsukiji	例文検索時に、ユーザの単語が複数出現した場合に重み付け ドキュメントの種類(wordやpptとか)毎に重み付けの係数の値を変えることで、学習するときの優先度を定める	2017-11-28 14:59:15

確認用ページ

結果

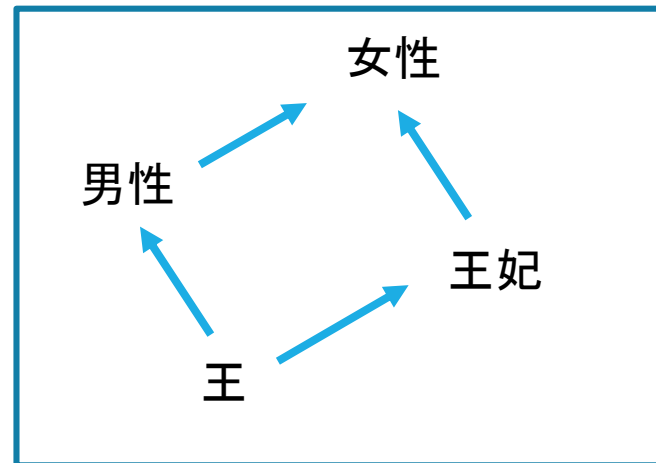
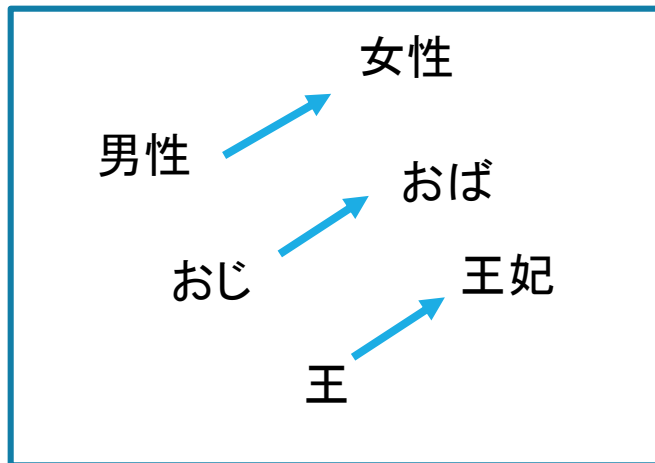
発言者	発言内容	タグ	発言元	発言時間	追加者
kurihara	Dolphinの環境を日本語に設定しました	解決 Server	2017-sotsuken	2017-11-20 15:47:31	ishikawa ishikawa
takano	あと、パスポートを取得していなくても申し込みはできます。	解決	general	2017-11-20 16:50:00	ishikawa
takano	あと、パスポートを取得していなくても申し込みはできます。	未解決	general	2017-11-20 16:50:00	guest
kaiseiaoki	:point_up:リマインダーを設定しました	解決	chat	2017-11-19 21:51:14	ishikawa
kaiseiaoki	技術的な資料については、研究や仕事などで英語での説明が求められる際などに発話するかと思います	解決	2017-sotsuken	2017-11-27 15:43:38	ishikawa
tsukiji	例文検索時に、ユーザの単語が複数出現した場合に重み付け ドキュメントの種類(wordやpptとか)毎に重み付けの係数の値を変えることで、学習するときの優先度を定める	解決	2017-sotsuken	2017-11-28 14:59:15	ishikawa

[他の発言を記録する](#)

記録一覧ページ

Doc2Vecで関連度抽出

- (前提技術) Word2Vecについて
 - 単語をベクトル (数値配列) で表現する技術のこと
 - 距離が近い単語は意味が似ていて、遠いと似ていない

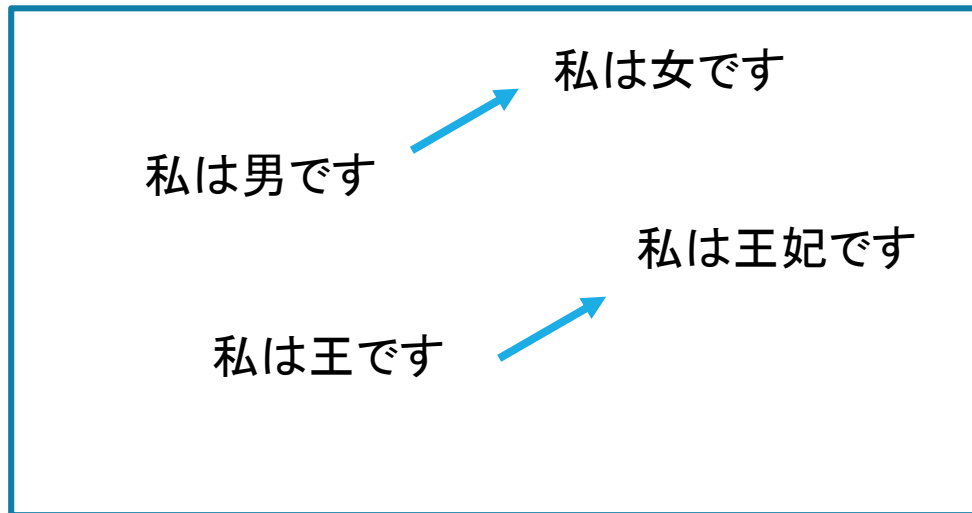


- 例1) $\text{王} - \text{男性} = \text{権力のある人}$
 $\text{権力のある人} + \text{女性} = \text{王妃}$
例2) $(\text{良い} + \text{最高}) / 2 = \text{より良い}$

Doc2Vecで関連度抽出

- **Doc2Vec**について

- Word2Vecの技術と同様に、文書をベクトル（パラグラフベクトル）で表現する技術
- 可変長の文章や段落に対応するので文章と文章の関連度の抽出が出来る



- 文章の要約やニュース記事の類似度抽出などといった活用法がある

検証

- 以下の文章を分かち書きした上で日常的な文章10個を含めて関連度を検証
 - 「php.iniの設定を変更することで、10秒以上の音声も録音することができました。」
 - 「php.iniファイルの変更が必要です」
- 一番関連している文章として抽出する事が出来たものの、値としては低い (0.29…).
- 何度か試行すると別の結果が出てくる事もあり、短い文章だと有効性があまり示せず、チャットの場合は短い文章で会話が続く事もあるのでDoc2Vecだと上手く抽出出来ないかもしれない事が分かりました.

実験概要

- 実際の発言からシステムに記録された情報が、振り返りに有用なのか確かめるための実験をおこなう.
- 「とあるシステムに新しい機能を追加するかどうかの議論（1対1）」を録音したものと、この議論をチャット上で再現したものを用意し、実験参加者を従来手法で記録してもらう方と本手法で記録してもらう2パターンに分けて比較する.

実験手順

1. 録音された議論を聞きながらメモを取る人（グループA）とチャット上に再現された発言を読んでいき、本システムに記録する人（グループB）に分ける.
2. 録音された議論を聴いてもらいながら、必要であればメモを取ってもらう（グループAの人）。チャット上に再現された議論を見てもらいながら必要であればシステムに記録してもらう（グループB）
3. 両方のグループとも用意された設問に回答してもらう.
4. 回答データを比較し、どちらが議論から多くの情報を取ってこれているかを考察する.

設問内容

- 「今回の議論で追加するのはどんな機能か？」
- 「この機能を実装する事は最終的に決定したか？」
- 「新しい機能を追加しようと発言したのは誰か？」

→以上の様な設問を両グループにおこなうことで、
議論で決まった事を後からでも確認出来る様な状態かを考察する。

今後の課題

- 現在と過去を結びつける良いタグが見つかるとは限らない.
- 音声からのラベル付けは音声認識の研究になってしまう.
- タグ毎の投稿数から「重要度」を取るのか, タグを付けた人から「重要度」を取るのかが定まっていない.
- 従来の手法の設定と提案方式の比較方法について
 - 個人が記録するメモなどでは議論で決まった事のみがメモされるのが殆ど.
 - 提案方式によってその決まるまでに至ったプロセスを示す事が出来る(?)