

# 技術系チャットを利用したブレインストーミングを対象とした情報統合

---

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あり、複数の手段を利用して頻繁にコミュニケーションがおこなわれている。
- 開発に関する情報が、個人でのメモや開発用ツールに蓄積されているものの、会話や議論まではプロジェクト全体に行き渡らないという問題がある。



# 関連研究（１）

---

## 【チャット・議論支援に関する研究】

- 『リアルタイムなコミュニケーション行為であるチャットへの意味タグ付加と電子ゼミナールへの適用』（2006,情報処理学会論文誌）
  - 発言に意味タグを付加した上でやり取りする機能を含むチャットシステムの提案
- 『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）
  - 対面式会議を総括的に支援する知識活動システムの開発
- 『Chatplexer：チャットを併用する口頭発表における発表者のための重要発言選択支援の試み』（2012,情報処理学会論文誌）
  - 口頭発表をした際の聴衆のチャットフィードバックから重要な発言の発見を支援するシステムの提案

# 関連研究（２）

---

## 【議論・議論参加者の分類】

- 『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)
  - 掲示板での議論から専門性やコミュニケーション能力を推定
- 『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)
  - 議論における6種類の役割を利用し、オンライン議論の分析
- 『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)
  - 係り受け構造と機械学習アルゴリズムによる分類・学習で議論理解を支援

# 関連研究（３）

---

## 【ブレインストーミング】

- 『あいづち機能を用いた分散ブレインストーミング支援システム』（２０１０,情報処理学会 情報処理学会論文誌)
  - 非対面型ブレインストーミングにおける相槌機能を利用したアイデア創出の提案
- 『ブレインストーミングにおける発話の可視化の影響に関する探索的な研究』（２０１６,人工知能学会 知識流通ネットワーク研究会)
  - 発話情報や発話数を表示する事で参加者の発言を促すシステムの提案
- 『なるほどボタン：褒める効果音ボタンを用いたブレインストーミング支援システムの検討』（２０１６,情報処理学会 エンターテイメントコンピューティング研究会)
  - 物理ボタンと効果音によるブレインストーミングの支援システムの考案

# 研究動機

---

- ソフトウェア開発に関わった時に「なぜこの技術なのか？」  
「この人に実装を任せれば早く終わっていた」といったコミュニケーションの不足が原因となっていそうな事が多くあった。
- 円滑なコミュニケーションが出来ていれば、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。

# 研究課題

---

1. 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.
2. 発言・議論内容の関連度をどのように抽出するか.
3. 過去の良い意見・アイデアをどのように現在の議論に統合するか.

# 研究目的

---

- 過去にあった良い意見やアイデアというものの自体を見つけやすくしたい.
- 過去にあった良い意見やアイデアを見つけ、現在の議論に紐付けることで、新たな良い意見やアイデアの創出に繋げたい.



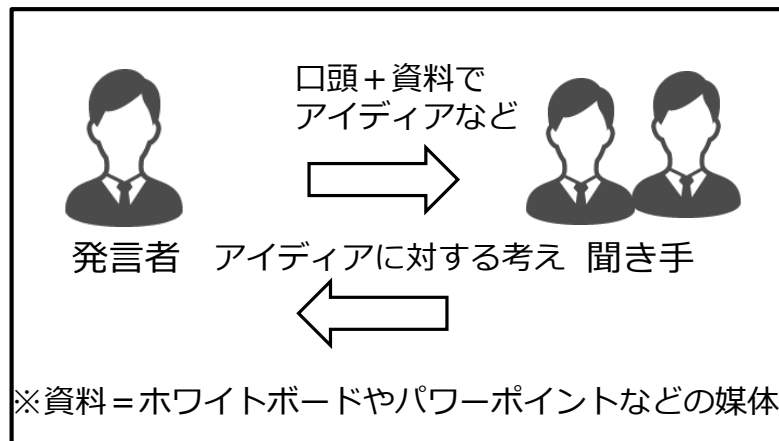
# 研究のアプローチ

---

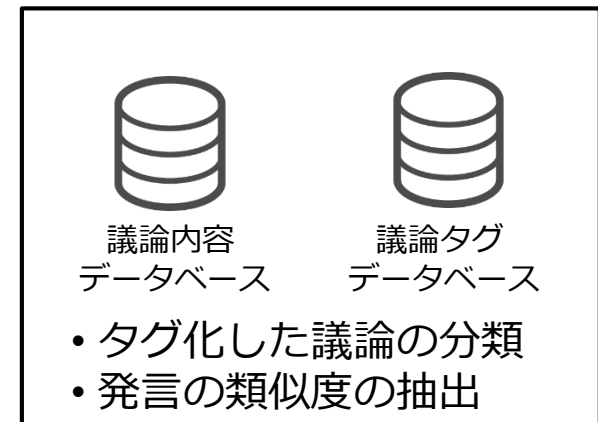
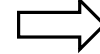
- ソフトウェア開発に関する議論やアイデアが出る場所として「**技術系チャットでのブレインストーミング**」を利用する.
- ブレインストーミング参加者に、チャット上で出てきたアイデアや意見をタグで記録してもらう.
- ブレインストーミングで記録されたアイデアや意見をスコア化して抽出することで、議論やアイデアの統合を目指す.

# 提案システム図

## ブレインストーミング

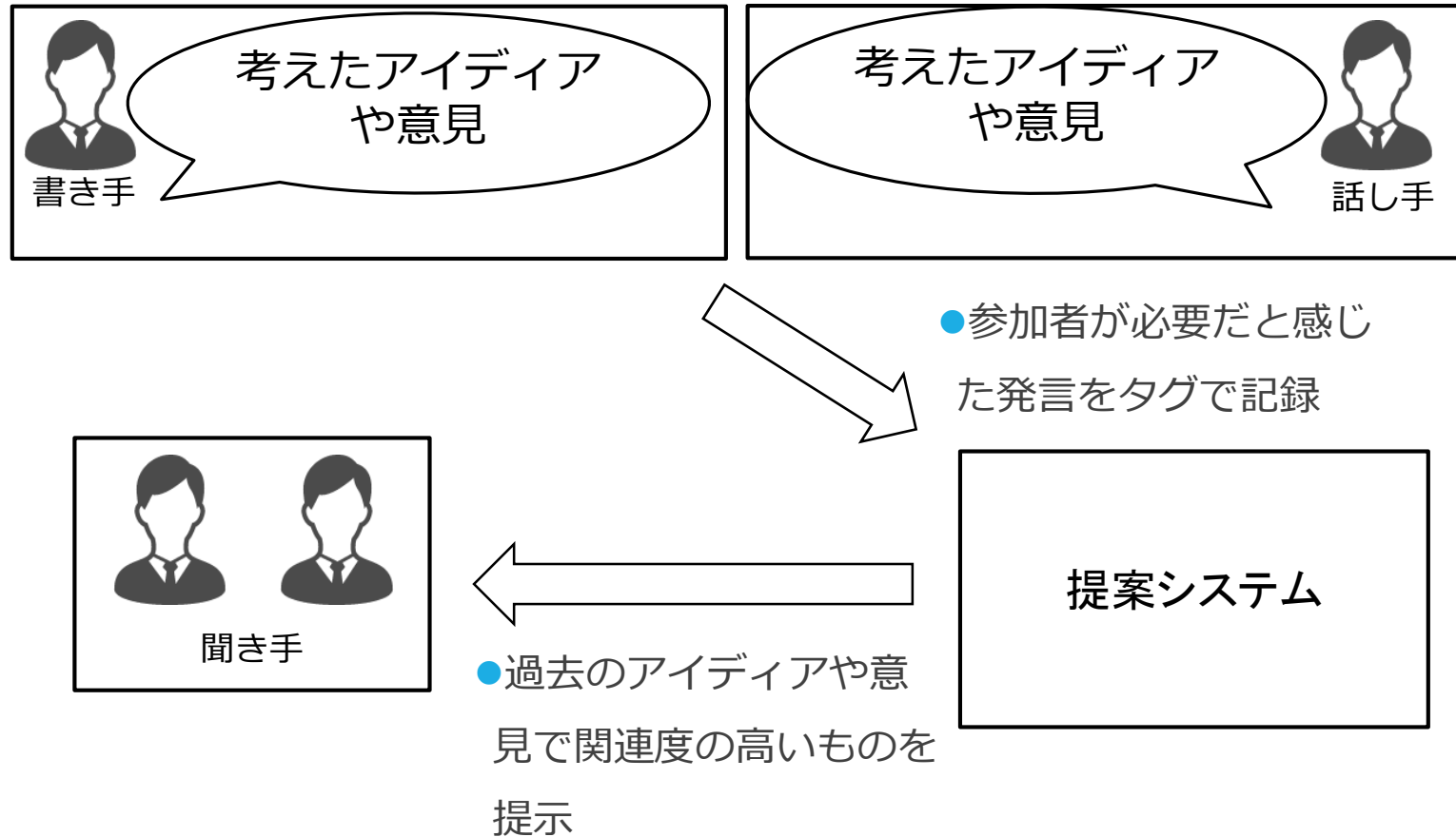


発言を記録



記録された情報  
をオンラインで  
共有

# 提案システムの流れ



# 従来方式との比較

---

## 【対面型BSの場合】

- テーマが違う所からアイデアが再利用されることは無い.

## 【非対面BSと本提案方式を利用した特徴】

- 過去のブレインストーミングの内容が比較される事で, アイデアが再利用される可能性がある.
- 過去の意見と統合することで, その場での議論では発見出来なかった考えやアイデアの導出に繋がられる.

# 実装

---

# 実装（１） 発言の記録

- 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

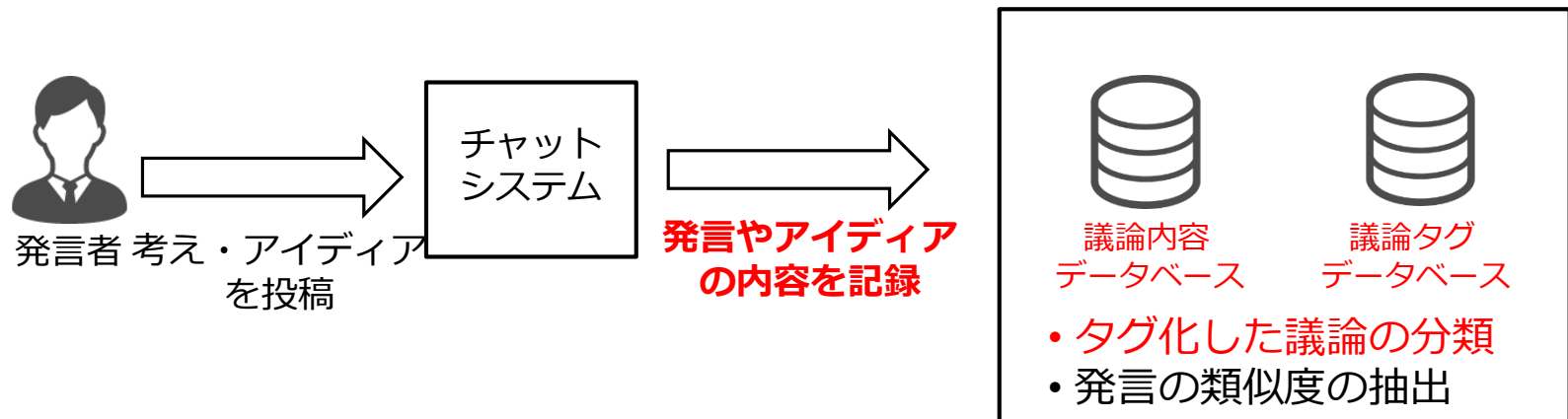
documents table

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）

---



# 実装 (3)

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

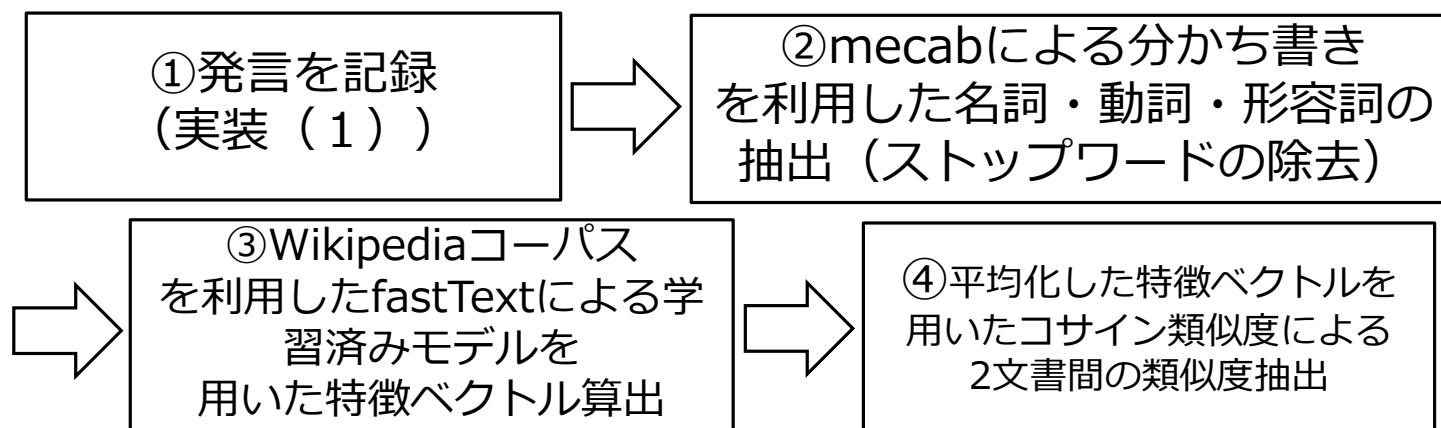
[他の発言を記録する](#)

## 記録一覧ページ



# 実装（４）

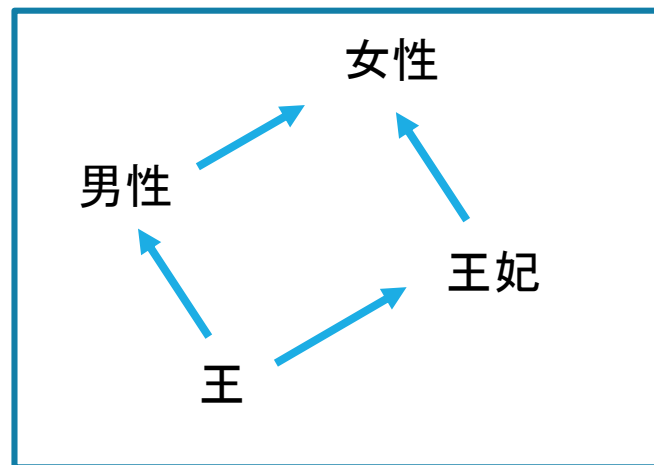
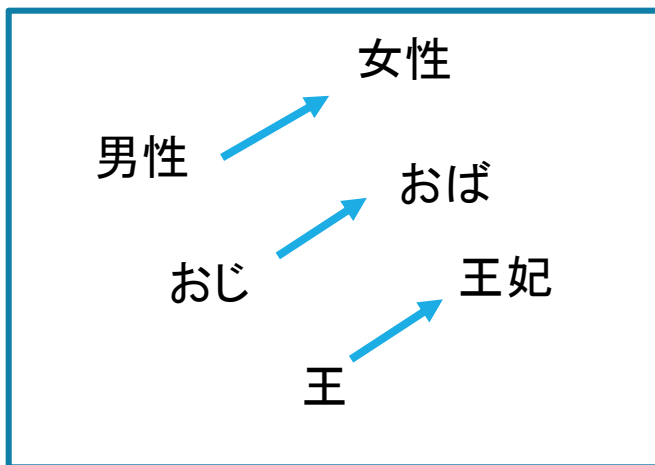
- 発言の類似度を抽出する機能の作成



- コーパスに存在しない語は特徴ベクトルを0とする.
- 学習済みモデルはfastTextと呼ばれるWord2Vecを応用した技術を利用して作成されたものを利用した.

# Word2Vec

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

# fastText

---

- Word2Vecと同様に単語の分散表現を獲得可能
- Word2Vecでは出来なかった活用形の考慮を可能にしている

例) Word2Vecでは「go」「goes」「going」は人間的には近いがモデル的には別の単語として認識

fastTextでは単語を構成する文字のまとまり(sub word)を考慮することで活用形で変化しない基幹部分の表現を共有可能(「go」と「goes」のgoは同じとして精度を高められる)

- 今回はWikipediaのテキストコーパスをfastTextを用いて学習された公開モデルを利用した

# 特徴ベクトルの生成

---

1. Wikipediaにある34万単語を**Skip-gram**でモデル化することで、特徴ベクトルを生成するSkip-gram：単語の周辺にどの単語が登場するかの確率でモデル化する手法
2. Wikipediaにある約34万単語が300次元のベクトルで示された学習済みモデルを読み込む
3. 用意した発言セットを分かち書きし、ストップワードを除去した上で学習済みモデルを利用して単語の特徴ベクトルを抽出する
4. 発言文章に登場する単語の特徴ベクトルの平均を計算
5. 平均化された2つの特徴ベクトルをコサイン類似度を利用して類似度を示す

# 実験

---

# 実験1 概要

---

- ソフトウェア開発のブレインストーミングで登場しそうな**発言**とそれに応答する発言を組み合わせた**発言セット**を用いて類似度を抽出する.
- 発言** : Slack上に利用者が投稿した1文章のこと.
- 発言セット** : Aが何かを発した事に対するBの応答をあわせたもの.

例) A『新しい機能ですがなんかありますかね?』

B『うーん、新しい機能よりも今ある機能で改善したいのがあります。』

## AとBの発言セット

『新しい機能ですがなんかありますかね? うーん、新しい機能よりも今ある機能で改善したいのがあります。』

# 実験1 目的

---

- ブレインストーミングで生まれるアイデアや考えが含まれる発言同士の類似度を示す事が出来るのかを確認する.

# 実験1 手順

---

1. 実験に利用する発言セットを用意する
2. 発言セット同士を比較し類似度がどう出てくるかを示す
3. 類似度が示せたとして、どうしきい値を設定すればいいかを考察する



# 実験1 結果

---

# 実験2 概要

---

- 複数名によるチャット上での**ブレインストーミング**を実施し、必要だと感じた発言をプロトタイプシステムを通じて記録してもらう。そして、記録されたアイデアや議論をスコア化して比較する。

# 実験2 目的

---

- スコアを用意して、ブレインストーミングで生まれたアイデアや議論の類似度をランキング的に示せるか（統合できるか）を見る。

# 実験2 手順

---

1. 評価パラメータを用意し, パラメータを利用したスコア算出の式を用意する
2. 記録された発言セットをその式を利用して評価する
3. 評価された発言セットのランキング結果を考察する

# 実験2 評価式

---

以下の評価パラメータを用いてブレインストーミングの評価付けをおこなう。

- 初期のアイデア数
- 発言数
- チャット以外（オンラインホワイトボードなど）を利用したか
- 解決方法の導出（導出にいたった経緯で細分化）
  - 全員一致でテーマに対する答えが出た場合
  - 誰かの反対や疑問がありながらも答えが出た場合
  - 全員反対な答えなどの中からあいにく見つけてきた場合
  - 答えが見つからなかった場合

# 実験2 結果

---

# 今後の課題

---

- 様々な手段での議論・会話への対応
- スコアにより分類された発言データを用いたコーパスの作成による自動分類
- ブレストの実施数と良いプロダクトの相関実験