

# 全文検索で Redmineを さらに活用！

須藤功平

株式会社クリアコード



Redmine大阪 第17回勉強会  
2017-08-26



# よく活用されたRedmine

## 情報の宝庫

- 問題解決の経緯
- Wikiに整理したノウハウ



# さらに活用するために

- 必要なときに
- 必要な情報に
- 素早くアクセスしたい！



Powered by Rabbit 2.2.1



# 活用の壁

- よく活用されたRedmine
  - 大量の情報（ただし**玉石混合**）
- さらに活用するための課題
  - 大量→必要な情報を見つけにくい
  - 大量→探す時間が増える

# 課題例1：検索ノイズが多い

検索: Redmine

結果 (74)

Wikiページ (63) チケット (10) 更新履歴 (1)

-  **Wiki: テレワークを考える会-20170407** (1491265723)  
# テレワークを考える会-20170407 14:00 - 15:00 ... とができなくて〇〇で支障がでそうだから、クリアコードメンバーは口頭で共有した内容を朝会后にRedmine/Zuulipに書くのはどう? みたいな話でテレワークの実現を考えることを期待 ... とができなくて〇〇で支障がでそうだから、クリアコードメンバーは口頭で共有した内容を朝会后に  
2017/04/04 15:02
-  **リビジョン 22edc017 (misc): Add more newlines to pretty format on Redmine** (148965326)  
2017/03/16 18:08
-  **Wiki: テレワークふりかえり-20170227** (1488173023)  
# テレワークふりかえり-20170227 2月27日 (月) 13:30-1 ... 計画通りに仕事が進んでいるかどうかを自ら管理・評価できる 参考: <https://redmine.clear-code.com/projects/clear-code/wiki/NE5%...+18+12=81W> ## その他 14:22-14:22 \* なし  
2017/02/27 14:23
-  **Wiki: ふりかえり2017-02** (1487739943)  
# ふりかえり2017-02 2017-02-24 (金) 16:00- ... nager リリース。:cake: # 2016-12 geoip2\_c を作った。redmine.clear-code.com の絵文字対応 :rocket: # 2017-01 共 ... くとよいのでは。 \* ライセンスチェックは本当に厳しいのでんばれ  
2017/02/22 14:05
-  **Wiki: ふりかえり2017-02(結城)** (1487728406)  
# ふりかえり2017-02(結城) ## 目標 \* Groongaサポート ... 立場で得た知見を、自分が後進に接する時に活かす。 ## 成果 <https://redmine.clear-code.com/projects/groonga-task/wiki/2017-01-31> <https://redmine.clear-code.com/projects/groonga-task/wiki/2016-12-28> <https://redmine.clear-code.com/projects/groonga-task/wiki/20...> そろ1時間~4時間の枠内で形にするという事を意識するスタイルに切り替えていっている  
2017/02/22 10:53
-  **Wiki: MozillaのBugzillaへのパッチのsubmitのやり方** (1487665420)  
# MozillaのBugzillaへのパッチのsubmitのやり方 ## はじめに ... illaでの開発プロセスを一通り行ったのでまとめます。[前回の記事](<https://redmine.clear-code.com/projects/clear-code/wiki/Gecko...>) よりも前から開発が続いているソフトウェアの開発プロセスを体験してみるのはいかがでしょうか。  
2017/02/21 17:23
-  **Wiki: GeckoエンジンにmacOSでprintToFileの機能を実装してみた話** (1487581754)  
# GeckoエンジンにmacOSでprintToFileの機能を実装してみた話 # ... パッチがコミットされる前のコードで設定されているかどうかを見ます。 <https://redmine.clear-code.com/projects/clear-code/wiki/Mac0...> Bug番号がまだ6桁であり、周辺のコミットに紐づいたBug番号は7桁なので感慨深いですね。  
2017/02/20 18:09
-  **タスク #3464 (新規: Redmineの検索精度向上)** (1487203025)  
Mrongaを導入して速くなったんだけど欲しいものを見つけれない  
2017/02/16 09:12

これ以外全部ノイズ



# 課題例2：検索が遅い

- 本当に大量だね！
- ↓これくらいだと遅くない
  - チケット数：4049
  - コメント数：41264
  - Wikiページ数：1081
  - 1秒以内に返ってくる
    - キャッシュが効く：0.1sくらい
    - キャッシュが効かない：0.7sくらい



# 解決策

Groongaで  
Redmineを  
高速・高精度  
全文検索



# 解決例1：検索ノイズがない

検索: **Redmine**

結果 (74)

Wikiページ (63) チケット (10) 更新履歴 (1)

 タスク #1795 (終了): 新しいVPSにサーバーを移行する (2861, 0)

OS は(まだ入っていないけれど) Debian wheezy にする。 \* [redmine.clear-code.com](#) \* [git.clear-code.com](#) (実体↑と同じ ... [it.clear-code.com](#) のディスク容量が足りないので移行する。 [git](#) と [redmine](#) は同時に移行する必要がある。 [git](#) の管理は [gitolite](#) じゃなくて [gitl](#) ... 既に使用しているソフトウェアについては Wiki の内容を参考にして作業する。 ([redmine.git](#)).clear-code.com のサーバーの切り替え自体は、休日にした方がいいかもしれない。  
2013/09/12 18:37

 タスク #3464 (新規): **Redmine**の検索精度向上 (804, 0)

[Mrroonga](#)を導入して速くなったんだけど欲しいものを見つけれないことが多い  
2017/02/16 09:12

全部それっぽい！

 タスク #1269 (終了): **Redmine**のWikiの変更通知メールにdiffを入れたい！ (520, 0)

h2. 問題 [Redmine](#)のWikiを更新すると変更通知メールがくるけど、そこには変更点が入っていないからいちいち[Redmine](#)を見にいかないといけない。これは、面倒くさいのでせつかくの変更を追いかけるのが億劫になって ... が入っていて簡単に変更点を追いかけるようになっている。 h2. 解決方法 [Redmine](#)本体を変更してパッチを送る？プラグインで実装する？  
2012/03/27 18:18

 タスク #3133 (終了): **Redmine**: コメント変更時にdiffを送れるようにする (401, 0)

朝会チケットとかで、とりあえず穴だらけでコメントを追加してその後で編集して完成、とされると編集分がdiffで流れてこないのが最新内容を把握するのがしんどい。  
2016/10/19 10:16

 タスク #951 (終了): さくらのVPSにrackspaceから移行する (160, 0)

h4. 現さくら \* [www.clear-code.com](#) : [squeeze](#) \*\* [ranguba](#) \* [redmine.clear-code.com](#) : [ubuntu](#) \*\* [redmine](#) \* [mail.clear-code.com](#) : [CentOS5](#) \*\* [militer](#) ... [ear-code.com](#) : [squeeze](#) \*\* [ranguba](#) (変更なし) \* [redmine.clear-code.com](#) : [ubuntu](#) \*\* [redmine](#) (変更なし) \* [mail.clear-code.com](#) : [CentOS5](#) \*\* ... [ruramal](#) (移行) [Ruby1.9](#)必須 \*\* [名札generator](#) (移行)  
2011/09/14 12:08

 Wiki: **Redmine** (137, 0)

# [Redmine](#) プロジェクト管理システム。 [[redmine.clear-code.com](#)] (<https://redmine.clear-code.com/>) で動いている。 ## アップグレード方法 作業は [redmine.clear-code.com](#) で行う。 基本的には [[Redmine](#) - [RedmineUpgrade](#) - [Redmine](#)] (<http://www.redmine.org/projects/redmine/wiki/RedmineUpgrade>) に書かれている方法に従う。 これを自動化したものが [redmine/bin/update.sh](#) に置いてある。 アップグレード方法が更新されることもあるので、新 ... グレードする前に念のためシェルスクリプトを確認すること。 なお、 2013-11-24時点



# 検索ノイズ？

- 上位N件でがっかりする？
  - がっかり→ノイズ多い
  - ユーザーは上位N件しか確認しない
  - Groongaはスコアを調整してソート
- 件数は関係ない
  - 実はこの例ではヒット件数は同じ



# 解決例2：検索が速い

- ↓これくらいだと誤差
  - チケット数：4049
  - コメント数：41264
  - Wikiページ数：1081
  - 参考：  
既存（キャッシュあり）と  
Groonga（キャッシュなし）が  
同じくらい



# 200万チケットのケース

<https://twitter.com/akahane92/status/733832496945594368>



**Kuniharu AKAHANE**

@akahane92



フォローする



200万チケット @MySQL でやってみたよ。検索時間は約380ms。 #Redmine の未来が広がって嬉しいな。ありがたいな。／Redmineで高速に全文検索する方法 - ククログ (2016-04-11) [@\\_clear\\_code](http://clear-code.com/blog/2016/4/11...)

7

リツイート

8

いいね



10:31 - 2016年5月21日

10:31 - 2016年5月21日



# ここまでのまとめ

- Redmineをさらに活用したい！
  - 素早く必要な情報にアクセス！
  - 検索をがんばらないと！
- Groongaで全文検索！
  - 検索ノイズが少ない！
  - 速い！



# ここからの話

---

導入したくなる  
情報を紹介



# インストール：RDBMS

- 対応RDBMS

- MySQLもPostgreSQLも！

- RDBMSにGroonga対応  
プラグインをインストール

プラグイン：別途サービス管理不要→運用コスト増加なし

- MySQL：Mroonga
  - PostgreSQL：PGroonga



# インストール：MrOonga

<http://mroonga.org/ja/docs/install.html>

- パッケージで簡単インストール
- 対応プラットフォーム
  - Debian GNU/Linux, Ubuntu
  - CentOS Fedora
  - macOS
  - Windows



# インストール : PGroonga

<https://pgroonga.github.io/ja/install/>

- パッケージで簡単インストール
- 対応プラットフォーム
  - Debian GNU/Linux, Ubuntu
  - CentOS Fedora
  - macOS
  - Windows

# インストール：プラグイン

```
% cd redmine
% git clone \
  https://github.com/okkez/redmine_full_text_search.git \
  plugins/full_text_search
% bin/rake redmine:plugins RAILS_ENV=production
```

再起動で完了！

# 高度な検索1：AND/OR/NOT

- キーワード単位でのAND/OR/NOT
  - 従来：AND/ORのみ
  - 従来：キーワード全体で切り替え
- 例：
  - (Groonga OR Mroonga) -PostgreSQL
  - GroongaまたはMroongaを含むがPostgreSQLは含まない

# 高度な検索2：並び順の変更

- ソート対象：スコア・更新時刻
  - 従来：更新時刻のみ
- ソート順：昇順・降順
  - 従来：降順のみ



# 並び順の変更UI

Redmine

☒ チケット ☒ ニュース ☒ 文書

▼ オプション

☐ 未完了のチケットのみ

☐ 添付ファイルと除外 ☐ 添付ファイル

検索結果の並び順

☒ スコア ☐ 更新日時

☐ 昇順 ☒ 降順



# ここまでのまとめ

- 簡単にインストールできる
  - 運用も楽
- AND/OR/NOTを使える
  - 絞り込みに便利
- 並び順をカスタマイズできる
  - スコアと更新時刻で十分



# 2017年5月以降の話

- redmine. tokyo 第12回勉強会
  - 2017年5月
  - ノイズが少なくて高速
- その後
  - 高速化
  - 類似issue検索対応



# 高速化

- 旧
  - LIKE検索をGroongaに置き換え
  - 他の条件はそのまま
- 新
  - すべてGroongaで検索
  - 高速・柔軟・高度な検索をサポート





# データの扱い

## ■ 旧

- 正規化してデータを保存
- 検索時に複数回のSELECT

## ■ 新

- 1つのテーブルにデータを保存
- 1回のSELECTで検索（**速い**）



# 柔軟

- ドリルダウンで対象を分別
  - チケットn件・Wikiページn件とか
  - 標準でもあるやつ
- ハイライト・スニペット
  - 標準よりもいい感じ
- 類似文書検索




# 類似文書検索

## 活用例

- 類似issue検索
  - 実装着手
- 類似Wiki検索
  - 未実装



# 類似issue検索






[groonga-dev, 04178] PGR0ONGAでレコード追加時にエラー発生 « 前 | 228/307 | 次 »



横山 昌史 が10ヶ月前に追加. 8ヶ月前に更新.

説明 引用

お世話になります。高見です。  
04175の件とは別件のエラーについての質問です。  
✦04175は諸般の事情で話が止まっており、現状では報告できることがありません。  
...

 pgroonga.zip (11.3 KB) groonga-dev, 04180に添付されていた Hayashi Kentaro, 2016/11/16 18:53  

類似チケット (43, 5ms)

|  |    |                 |  |
|--|----|-----------------|--|
| タスク#3245: [groonga-dev, 04194] PGR0ONGAでのデータ異常の検出 (11345400) | 終了 | 結城 洋志           |  |
| タスク #4071: PGR0ONGAのインデックスファイルの容量を取得する方法 (9618672)           | 解決 | Hayashi Kentaro |  |

子チケット 追加

関連するチケット 追加

関連するチケット 追加

# 類似 issue : ユースケース

- 新しい問題が発生！
- あれ？どこかで見たような…
- プラグイン：
  - 「このissue、似ていませんか？」
- →問題の早期解決！



# 類似issue検索：現実装

- issue閲覧時
  - 自動で類似issueを提示
- Groonga内のデータを活用
  - テキストデータで類似文書検索
- 精度
  - まあまあ



# ここからの話

---

## 今後の展望



# 類似issue検索精度向上 メタデータを活用

- 手動設定の関連issue情報使用
- 属性情報使用
  - カテゴリとかトラッカーとか
- 同一プロジェクトに限定する
- Wikiも対象に加える





# 類似issue検索精度向上 機械学習技術を活用

- 字面でなく意味で類似文書検索
  - 例：メールのthreadとプログラムのthreadは違う
  - 例：MroongaでMySQLも類似度向上
- 回答の自動生成
  - RNNを用いた機械翻訳とか
  - 類似issueの結論を探さなくてよい



# データ活用に向けて

- GroongaにPython組込
- Rubyでデータ処理
  - Red Data Tools
- Apache Arrow
  - 次世代データ交換フォーマット
  - コミッターになった！



# 類似issue 共同開発組織募集

- 精度向上にはデータが不可欠
- データを持っていそうな組織
  - 大量ユーザー・長期間運用組織
  - マネージドサービス提供組織
- Redmine内データを活用したい組織はぜひ！

# 添付ファイル内全文検索

- 現状：ファイル名と説明
- テキスト抽出できれば検索可能
  - テキスト抽出ツール例：ChupaText  
Groongaプロジェクトのテキスト抽出ツール

# 添付ファイル内全文検索 実装案

- テキスト抽出部分をプラグインできるようにすればRedmine本体に入れられるのでは！？
  - ChupaTextよりApache Tikaを使いたい人多いそう
  - 検索はLIKEでよい  
Groongaプラグインを入れたら高速になる、でよい
- だれかチャレンジしない？



# 入力補完

- 検索ボックス
  - 入力に応じてキーワードを随時提示
  - ↑ ヒットするもののみ
- 各種入力欄（タイトルなど）
  - 入力に応じて既存内容を随時提示
  - ローマ字でインクリメンタルサーチ

# 入力補完：ユースケース

- 検索ボックス
  - 自然と適切な検索に誘導！
  - 例：typoを防ぐ・表記揺れを吸収
- 各種入力欄
  - 自然と表記を統一！
  - →検索しやすい・理解しやすい
  - 例：typoを防ぐ・表記揺れを吸収



# 入力補完：実装

- Groonga内のデータを活用
  - どのくらいヒットするキーワード？
  - ログも活用：検索頻度は？
  - メタデータも活用：  
プロジェクト・トラッカー毎に  
文脈をわける
- 開発協力者を募集
  - 実装・データ提供・発注など





# 同義語展開

- 表記揺れを吸収
  - 例：「打ち合わせ」・「打合せ」
  - 吸収：「打ち合わせ OR 打合せ」
- 同義語は管理が大変！
  - 文脈で異なる！増減する！
  - 同義語管理を支援
  - 例：同義語候補を自動生成

# 同義語展開：ユースケース

- 検索ボックス
  - 欲しい情報にたどり着ける！
  - 表記揺れによる未発見を解消
- 同義語管理のコスト削減
  - Redmine内のデータをより低コストで有効活用！



# 同義語展開：実装

- Groongaの組み込み機能を活用
  - 高速に同義語展開が可能
- 同義語候補を機械学習
  - Groonga内のデータを活用
- 開発協力者を募集
  - 実装・データ提供・発注など



# スマートナビ（仮）

- 言わなくても欲しいものを提示
  - 例：駅の近くでスマホを見る  
→その駅の時刻表を**自動で**表示
  - ポイント：ユーザーは  
「駅名 時刻表」で  
**明示的に**検索して**いない**！

# スマートナビ:ユースケース

- .o0（打ち合わせのアジェンダをWikiにまとめておこう）
  - Redmine：ホーム画面に該当Wikiページのリンクを表示



# スマートナビ：実装

- 行動履歴から知識を学習
  - 打ち合わせ前にいつもWikiを更新
- 内容から事実を収集
  - Wikiページ内のテキスト  
「参加者：○○…」から抽出
- 開発協力者を募集
  - 実装・データ提供・発注など



# ここまでのまとめ

- このプラグインの現状
  - Redmineの不便を解消  
(マイナスを少しプラスに)
- このプラグインの今後
  - Redmineをさらに有効活用！  
(少しプラスをすごくプラスに)



# ここからの話

---

Redmineの開発に  
参加しよう！





# 開発に参加？

- Redmine本体のコードを書く  
だけじゃない
  - バグレポート・テスト・issue対応
  - ドキュメント作成・宣伝
- プラグインを作る  
だけじゃない
  - 本体と同様↑のことも大事な開発



# 開発の参加例

- Groongaプラグイン開発時の知見をフィードバック
  - Defect #21993: サイズ上限が小
  - Patch #23153: フックを追加
  - Feature #25198: スコアを表示



# 敷居が高そう？

- OSS Gateワークショップへ参加してみよう
  - 「OSS開発に参加」を経験できる  
(Redmineもプラグインの多くもOSS！)  
(もちろんこのプラグインもGroongaもOSS！)
  - <https://oss-gate.doorkeeper.jp/>

 添付ファイル内全文検索を  
実装したい！

---

私に相談だ！



# まとめ

- Redmineをさらに活用するならGroongaプラグインを使おう！
- Redmineの開発に参加しよう！
  - OSS Gateワークショップが役立つ
- 宣伝：全文検索のことはクリアコードに相談しよう！