

ASKサービス

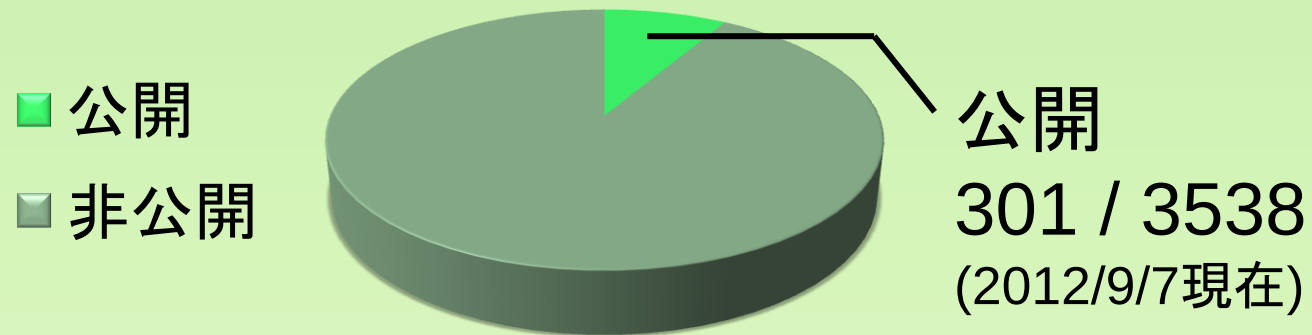
埋もれた事例を掘り起こす

2012年9月20日

参考調査係 岩井

背景

事例の大半が非公開



(今年度からの方針)

- ASK事例は、公開のものを除いて、5年で削除（個人情報管理の観点から）
- ただし公開用事例は残す

非公開の理由

- あまりにもFAQすぎる
 - 学外者ですが利用できますか？（同様多数）
- あまりにも個別的・一時的すぎる
 - 「〇〇」というジャーナルにアクセスできない
 - 単にそのジャーナルを契約してなかった
 - 単にそのジャーナルがメンテナンス中だった
- プライバシーに関わる
 - 自分の祖父の△△が東京帝国大学◇学部の卒業生だが、卒業年や学科を知りたい

これらは全く使えないのか？

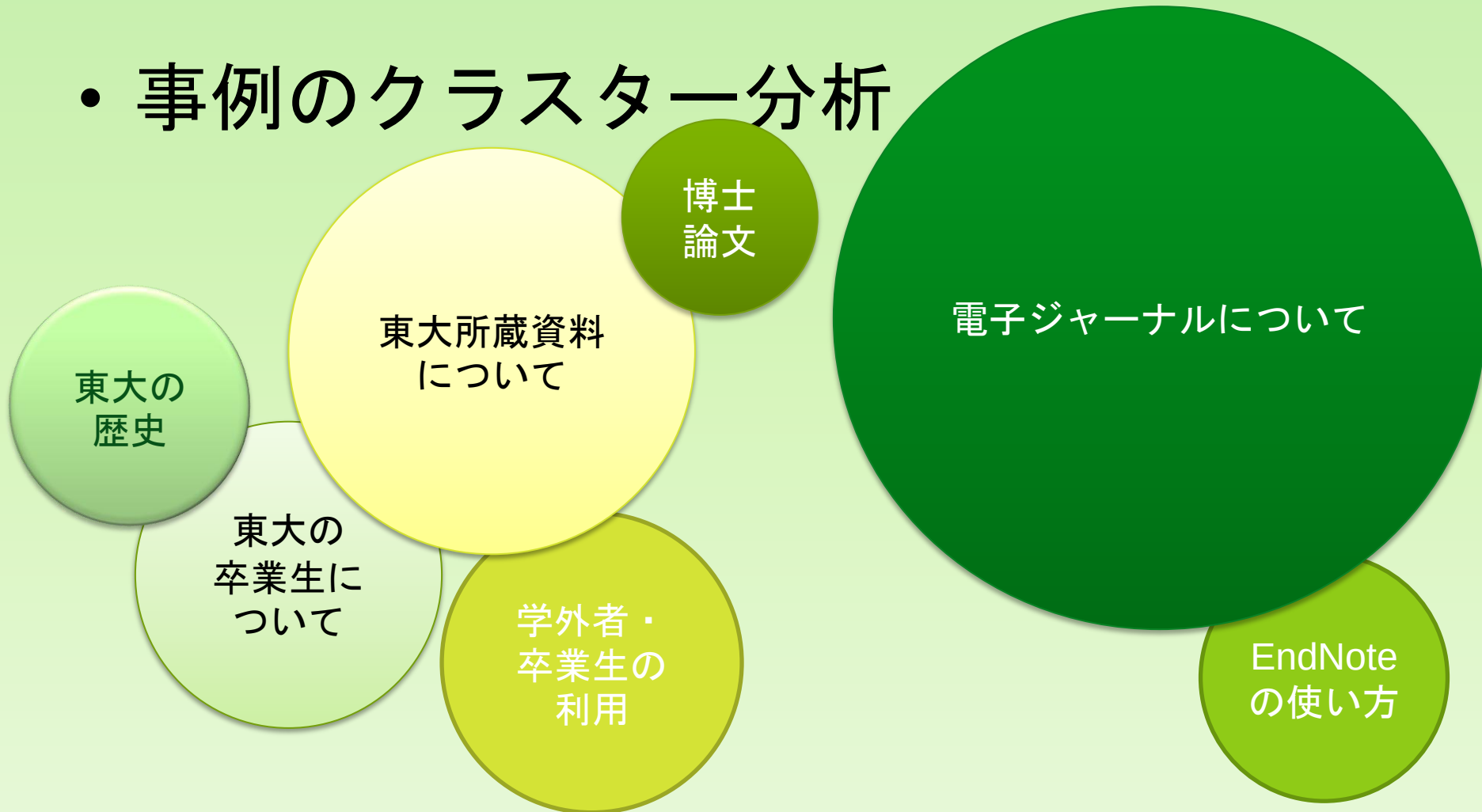
- FAQはFAQとして別の場所に載せておいたり、サイト改善の手がかりにしたりできる
- 個別的な質問も、固有名詞を除いた構造はFAQ になるのでは
 - 電子ジャーナルが見られないときは？
 - 卒業生について調べたいときは？

やりたいこと

- 多数のASK事例に共通する部分を浮かび上がらせることで、見過ごしていたFAQを可視化する
- ひとまずは既存事例に対して実施
- 理想としては、作業者が回答時に参照して、自分の扱う質問に近い事例を探せるツールに

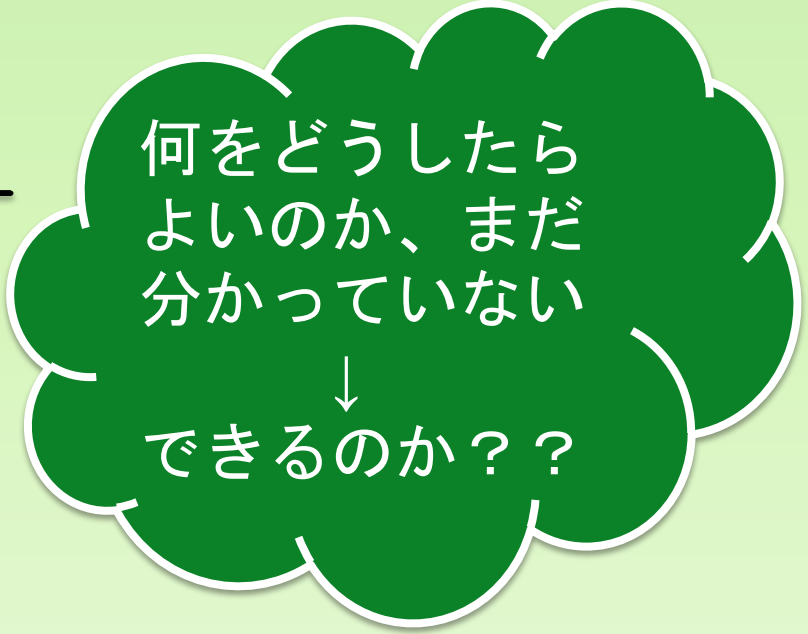
方法

• 事例のクラスター分析



必要なこと

- Ecatsからのデータ取得
- 自然言語処理
- クラスタリング
- 見せ方



何をどうしたら
よいのか、まだ
分かっていない

↓
できるのか??

参考：単語抽出の結果

抽出語	出現件数	抽出語	出現件数
図書館	2451	検索	1161
お願い	2320	図書	1138
利用	2013	所蔵	1110
資料	1679	ジャーナル	1090
東京大学	1582	契約	1018
総合	1512	閲覧	997
回答	1405	研究	956
情報	1230	問い合わせ	949
確認	1208	申し訳	914
電子	1163	論文	906

(使用したツール KH Coder <http://khc.sourceforge.net/>)

参考：品詞別の出現件数

名詞	件数	サ変名詞	件数	地名	件数	組織名	件数
図書館	2451	お願い	2320	柏	122	東京大学	1582
資料	1679	利用	2013	本郷	112	東大	488
情報	1230	総合	1512	駒場	101	国会図書館	174
電子	1163	回答	1405	アメリカ	59	医科研	39
図書	1138	確認	1208	目黒	47	生産技術研究所	35
ジャーナル	1090	検索	1161	文京	39	朝日新聞	28
論文	906	所蔵	1110	中国	38	国文学研究資料館	22
雑誌	702	契約	1018	江戸	28	東北大学	18
下記	688	閲覧	997	韓国	28	日経	17
方法	573	研究	956	米国	27	京都大学	16