

文書オントロジーを利用した 文書検索システムの研究

2011.7.13

井口 勝

東京農工大学 工学府 電子情報工学専攻

小谷 善行

東京農工大学 工学研究院 先端情報科学部門

人口知能学会「知識・技術・技能の伝承支援研究会」
SIG-KST-2011-03-04(2011-03-09)より抜粋・一部追記

内容

- 1. はじめに
- 2. 文書管理体制とその問題点
- 3. 提案システムの概要
- 4. 文書検索オントロジーの概要
- 5. 文書チャンクの概要
- 6. 今後の予定

1. はじめに

□ 知識の伝承

方法
人 → 人



Tool(知識伝承の教科書)
= 文書(電子、紙)



2. 文書管理体制とその問題点

- 文書は、企業経営・組織の形態や社会環境変化に伴って、その管理方法が変わることがある。



- 文書の同一方法での永続管理は、非常に困難。

2. 文書管理体制とその問題点

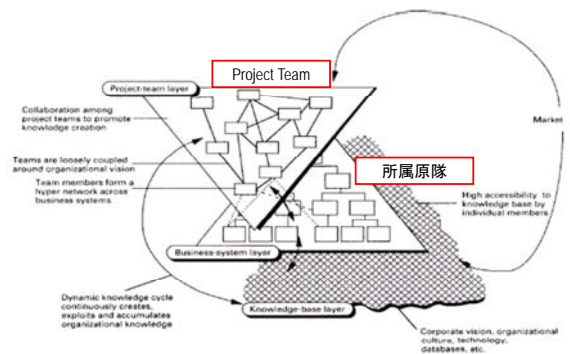
- 文書作成の容易化に伴う、文書の量増加。
- 組織改編、他社との経営統合などによる文書管理体系の変化。

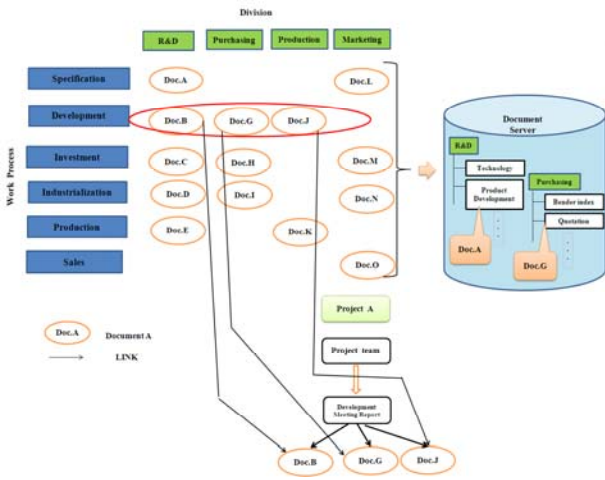


- 文書内容の細分化→文書管理番号の細分化→ 管理の困難性増加。

3. 企業組織の変化

Organization of Hyper Text Type





3. 提案システムの概要

①あるドメイン(企業、組織)における、
大量の文書に対する検索容易化。

↓
検索者の思いを機械化する。

↓
検索オントロジーの構築

3. 提案システムの概要

②あるドメイン(企業、組織)における、
大量の文書のクラス分け

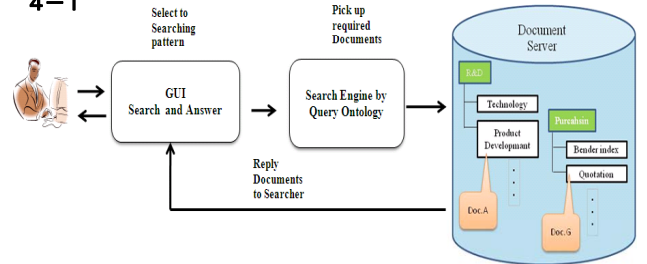
↓
細分化が過ぎると、管理が困難。

↓
ISO9001に規定した管理体系+
文書の意味チャックによる分類

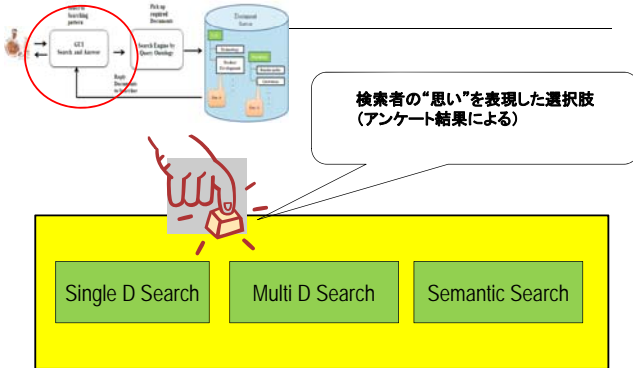
結果(文書)を、
推奨→機械
選別→人

4. 提案システムの概要

4-1



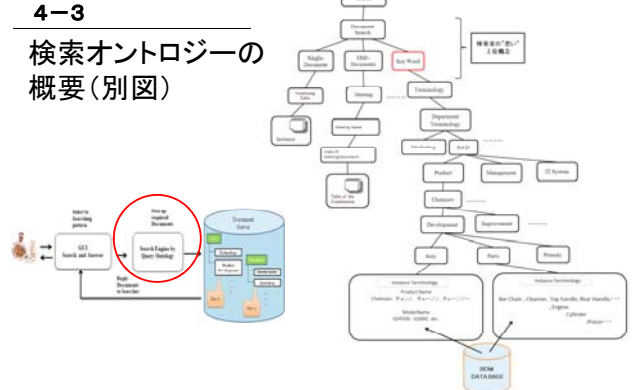
4. 提案システムの概要 4-2



4. 提案システムの概要

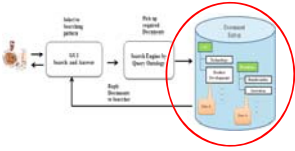
4-3

検索オントロジーの
概要(別図)



4. 提案システムの概要

4-4

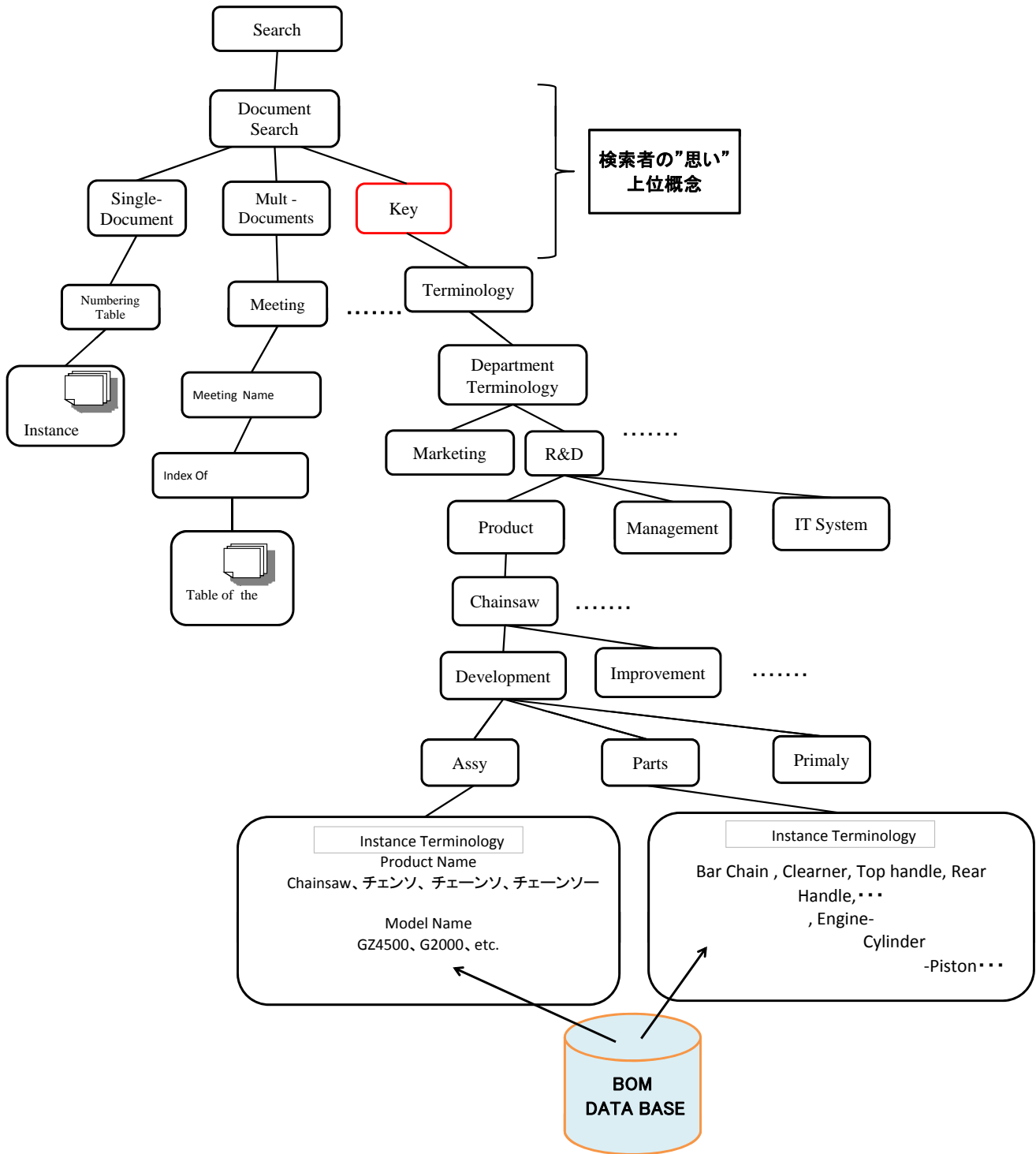


文書検索とチャンクの概要(別図)

5. 今後の予定

- ①検索オントロジーの下位概念を構築する。
 - ・製品、機種名、部品名、「材料名」
- ②文書のチャンクを実行する。
 - ・ネックは文字数の少ない文書、助詞のない文書、画像データ等の処理方法。
- ③全文検索との比較(有用性の確認)

以上

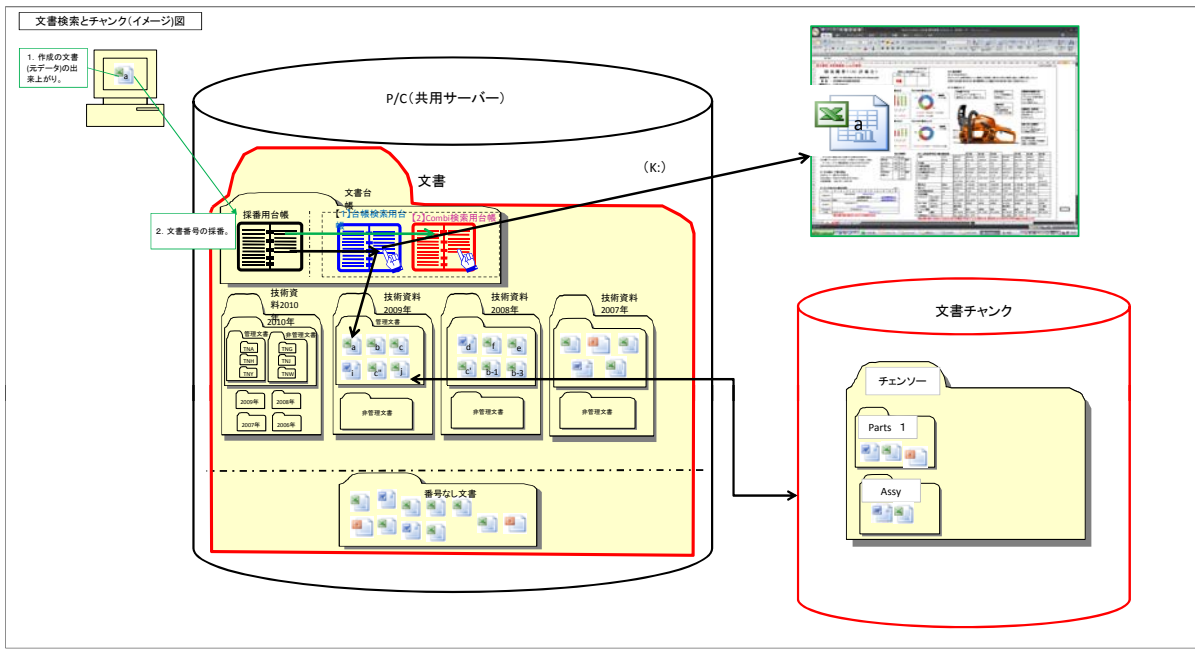


文書管理の模式図

※文書=各々が作成した技術資料(管理文書、非管理文書を含む)

2011年5月18日、現在
ハスクバーナ・ゼノ株式会社
開発センタ 企画管理G

緑線による文字、線、矢印の表記について、今後の管理文書の流れのイメージです。



【1】文書台帳検索 ~image~

Excel文書台帳からリンクで文書の '元データ' と 'SCANデータ' を開ける検索システム。

