

# 原子カプラントラブル報告書文書 からのテキストマイニング

東北大学大学院 技術社会システム専攻  
高橋 信

# 航空システムの安全性向上



航空管制システム

人間・機械系シミュレーションによるシステム評価

# 先進原子力システムの安全性に関する研究

- プルトニウム利用の安全性
- 放射性廃棄物の分離変換の安全性



リスク評価・管理

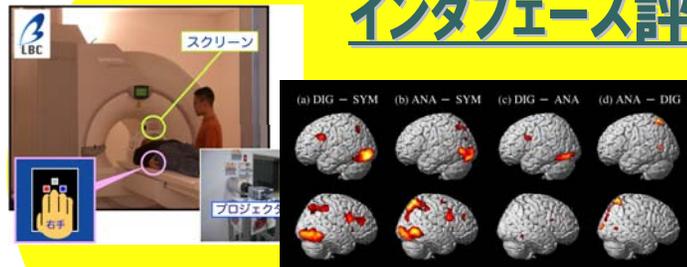
# 大規模システムの安全性向上

## 訓練用異常シナリオの分類と評価



原子炉運転訓練の効率的なプログラムの構築

## 脳機能解析を用いたインタフェース評価



脳機能解析手法のインタフェース評価への応用

# ●●● | 情報共有化の例

## インターネット

### ➡ ニューシア (<http://www.nucia.jp/>)

原子力施設情報公開ライブラリー 平成15年10月1日

発電所別の機器の故障率、信頼性



### ➡ JNES (<http://www2.jnes.go.jp/atomdb/jp/index.html>)

原子力安全基盤機構 平成15年10月1日

トラブル報告書を公開



原子力に関する情報が蓄積

# ●●● | 失敗知識データベース

## 文部科学省

平成12年8月より失敗知識活用研究会を開催し、**失敗情報の知識化、共有化及びその活用**の在り方について検討



## 失敗知識データベース

2005年3月23日より



科学技術分野の事故や失敗を未然に防止し、技術の信頼性と社会の安全性の向上

事例の分析で得られる教訓を共有できる知識として整理し、失敗知識データベースとして公開・普及

# 背景

## 事例知識活用という観点からみた場合

### ◆ 情報の共有化

トラブル報告書の蓄積のみ

### ◆ 失敗知識データベース

原子力分野:28件

- ・事例の網羅性が達成されていない
- ・人による主観的な見方や対象の知識
- ・結果が変動

## 事例を網羅し一貫性のある情報抽出方法



解析可能な事例表現(数値表現)によって  
統一的な基準を確立し事例を網羅

# 最終的なアウトプットイメージ

## 原子力プラント経年劣化事象



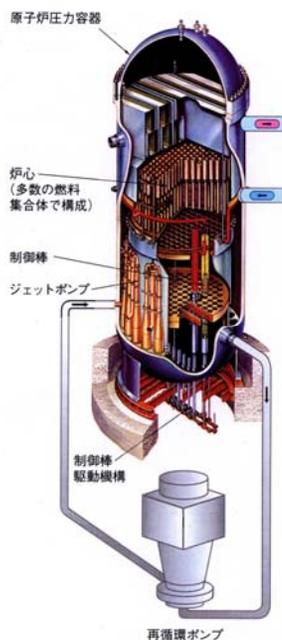
### 他産業分野での劣化事象



インターネット、文献内の知識



## データマイニングによる統合的な知識抽出



•設計変更時（材質、環境条件）に、新たな劣化事象発生の可能性の検討を支援

•他産業分野での経験からの外挿を考えた場合、原子力プラントでの劣化事象発生の可能性の検討を支援

→**プロアクティブ対応支援**

•新たに劣化事象が発生したときに、その事象が既知の発生メカニズムで説明可能か否か

→**未知事象の早期認識の支援**

**発生した後になって「予見できていたはずだった」という事象の発生を極力抑えることが最終的な目的**

## 背景

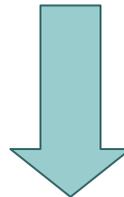
トラブル報告書

— 詳細度・書式の相違 —

事例数の増加

— 統一的な解析が困難 —

大量のデータから統一的な基準によって解析を行い、尚且つ事例を網羅

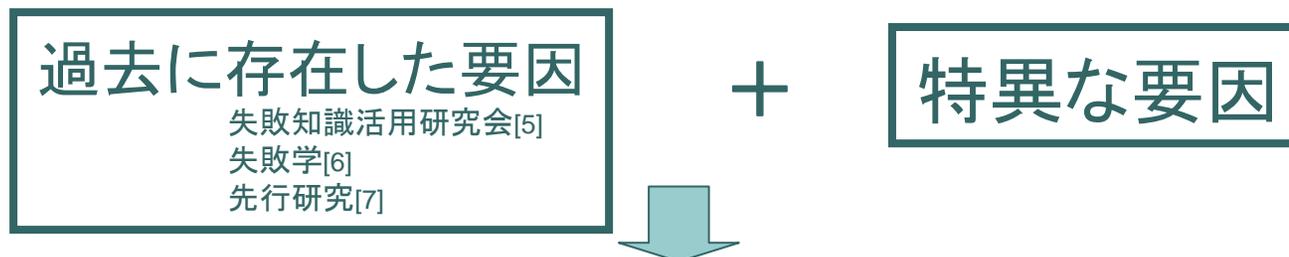
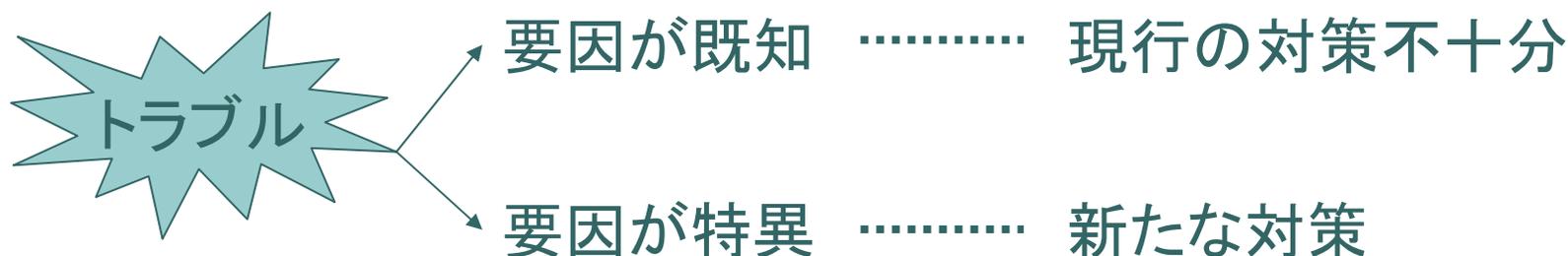


**テキストマイニング技術に着目**

# ●●● | 背景

➡ テキストマイニング技術を用いて事例特異性を抽出

特異な事例をただの一過性の事例としてとらえるべきでない

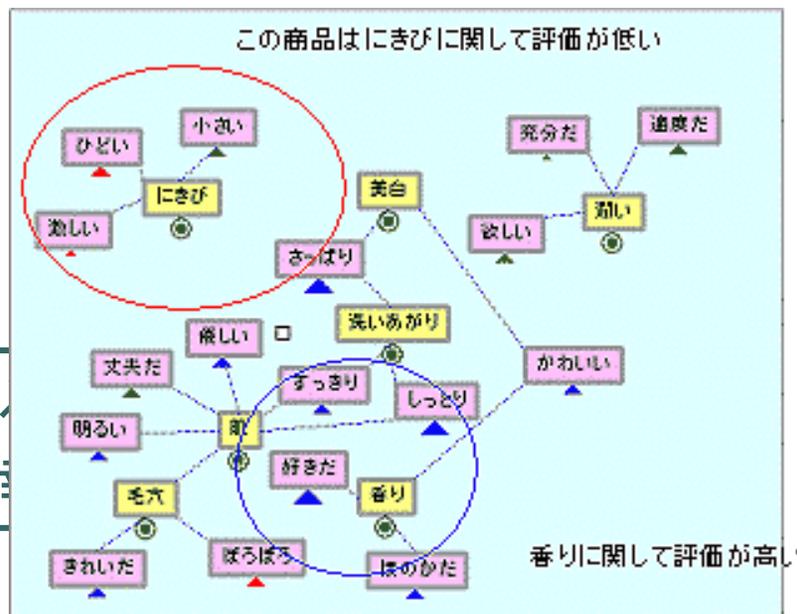
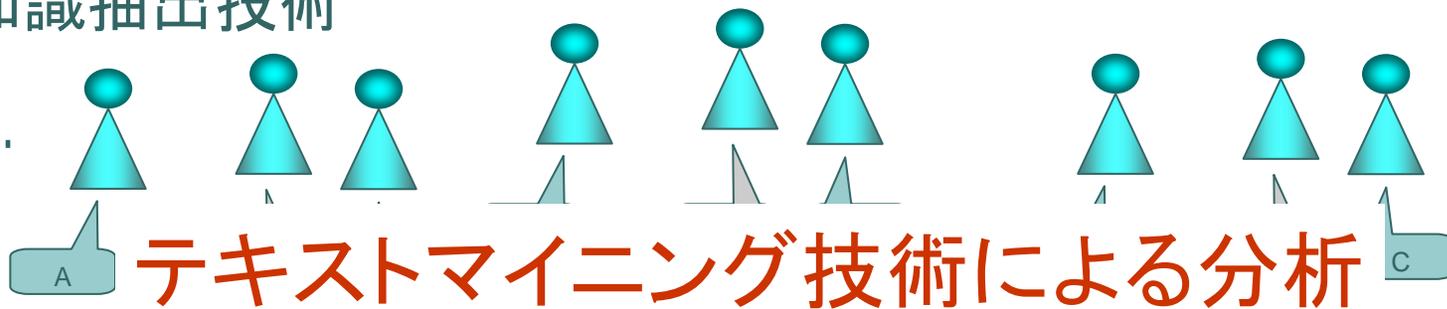


より網羅的な事故・トラブルの未然防止

# テキストマイニング技術

## 概要

大量の**テキストデータ(電子データ)**を用いたコンピュータによる知識抽出技術



ある化粧品に対する  
意見から得た情報

# テキストマイニング技術に関する 先行研究

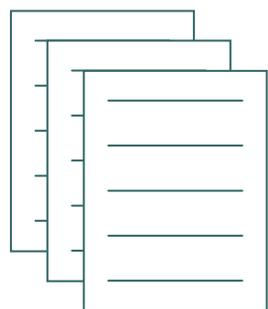
## 故障生起知識の構造化のための文書情報処理に関する研究

2002年度 内松

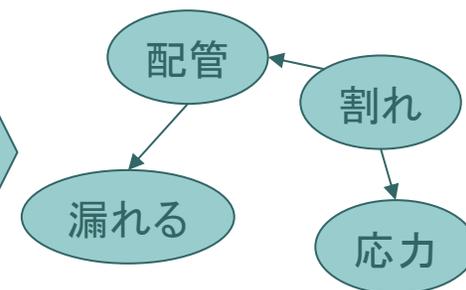
事件事例報告書の内容を近似的に単語と単語群の関連度を要素とした行列で表現する。関連度を算出する基準は、文章内での単語間の距離による重み付けを使用する。

行列表現された事例を用いて事例知識抽出を行う。これは、行列の類似度をコサイン尺度によって求め、事例の特異性を算出するものである。

事件事例文書



$$\begin{Bmatrix} \text{co}(1,1) \cdots \text{co}(1,m) \\ \vdots \\ \text{co}(m,1) \cdots \text{co}(m,m) \end{Bmatrix}$$



# ●●● | 目的

## 原子力発電所のトラブル事例報告書からの失敗事例知識の蓄積及びその抽出

### ▶ 事例群の解析可能な表現

多数の事例に適用可能な事例報告書の内容を単語間の関連度を要素とする行列で表現する。

### ▶ 事例知識の抽出

本研究では事例知識の中で特異性に注目し、数量的に抽出する。特異な事例の発見で普段気づかないプロセスの欠陥を見出し、想定外の事故・トラブルを軽減することができると思う。

# ●●● | 事例内容とは

文書中で内容を示すもの



単語  
特にキーワード

単語やキーワードの集合だけでは文書内容は明確にならない



キーワードとキーワードの関係の強さ(関連度)を算出

事例内容を表す行列を作成

# 事例の構造化

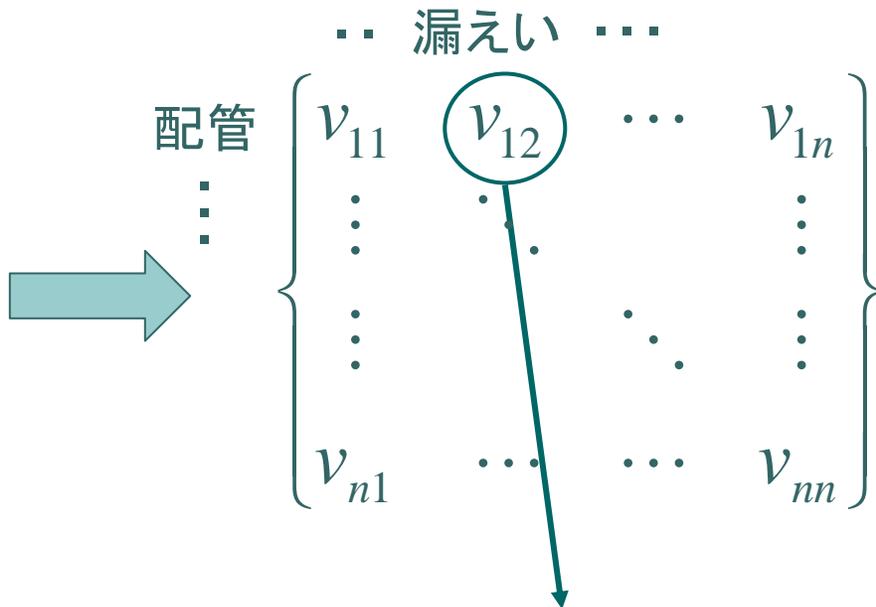
## 事例A

nは文書に現れる総単語数

原子炉給水ポンプ駆動タービンドレン配管管台部からの漏えい  
 中部電力浜岡原子力発電所2号喫沸騰水型、定格出力84万ワット  
 (ト)は、定格出力にて運転中のところ、11月3日午後9時10分頃、  
 タービン建屋3階にあるタービン駆動原子炉給水ポンプ室の床に水  
 が溜まっていることが認められた。点検を行ったところ、高圧ター  
 ビンからの蒸気を使って同ポンプを駆動するタービンの小口径ドレ  
 ン配管(同タービンを駆動した蒸気中に含まれる水分を回収して復水

器へ移送するための配管)付近から漏えいしていることが確認された。このため、当該配管の詳細調査を実施することとし、  
 3日午後10時50分から出力降下を開始し、4日午前7時48分に原子炉を手動停止した。なお、外部に対する放射能の影響はない。中  
 部電力浜岡原子力発電所2号喫沸騰水型、定格出力84万ワット  
 (ト)は、定格出力にて運転中のところ、11月3日午後9時10分頃、  
 タービン建屋3階にあるタービン駆動原子炉給水ポンプ室の床に水  
 が溜まっていることが認められた。点検を行ったところ、高圧ター  
 ビンからの蒸気を使って同ポンプを駆動するタービンの小口径ドレ  
 ン配管(同タービンを駆動した蒸気中に含まれる水分を回収して復水  
 器へ移送するための配管)付近から漏えいしていることが確認され  
 た。このため、当該配管の詳細調査を実施することとし、3日午後1  
 0時50分から出力降下を開始し、4日午前7時48分に原子炉を手  
 動停止した。なお、外部に対する放射能の影響はない。調査の結果、  
 4本ある当該ドレン配管のうちの1本の管台部(配管が給水ポンプ  
 駆動タービンに接続されている部分)に、直径約5mmの貫通孔が  
 確認された。また、当該部及び他の小口径ドレン配管の管台部に、  
 浸食と腐食の繰返しによる

ル) 格納槽) 点検 1 対正 11 階内 0が の後 手果、 をた。 隙時そ

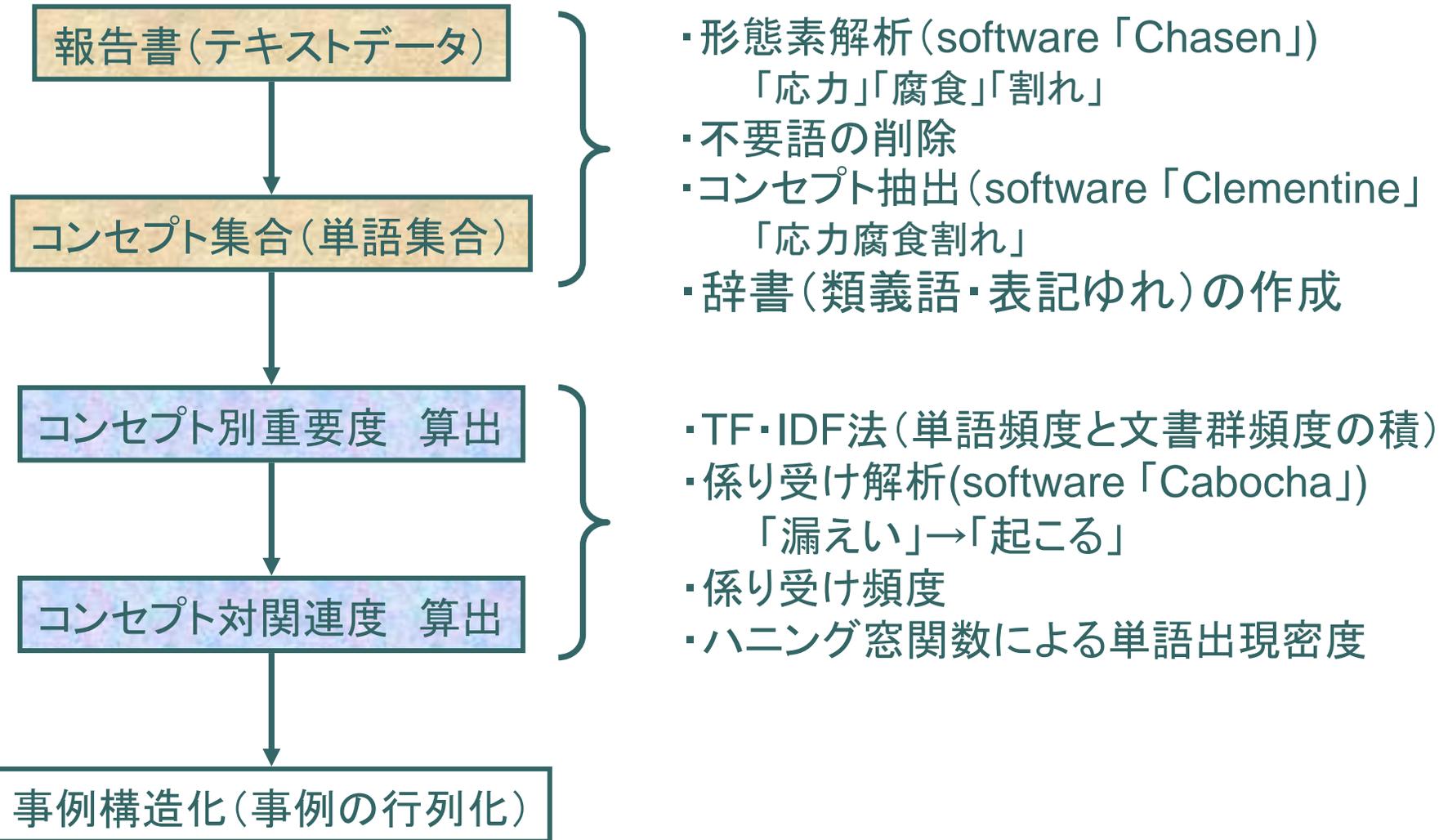


事例内容を単語の関連度で行列化

事例構造化

総事例数: 139事例

# ●●● | 事例構造化の全体像



# ●●● | 事例構造化

報告書(テキストデータ)

→ コンセプト集合(単語集合)

応力腐食割れがキャンピーシール部内面から漏えいに至ったものと推定される。



動作

応力 腐食 割れ が キャンピー シール 部 から 漏えい 至った もの に される と 推定

形態素解析

不要語の削除

# 事例構造化

報告書(テキストデータ)

→ コンセプト集合(単語集合)

応力腐食割れがキャンピーシール部内面から漏えいに至ったものと推定される。



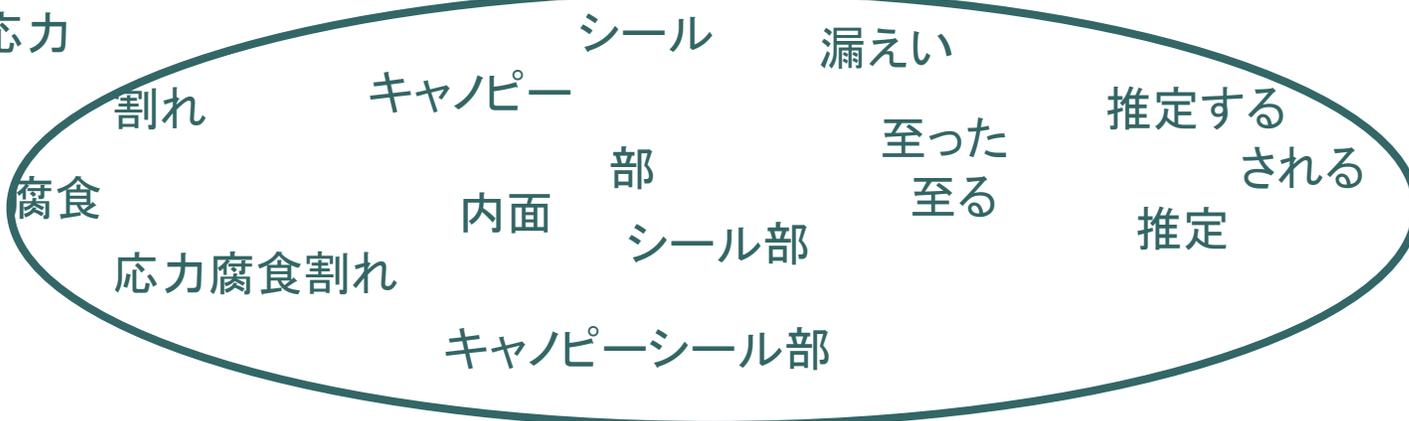
動作

形態素解析

不要語の削除

コンセプト抽出

辞書使用



コンセプト集合

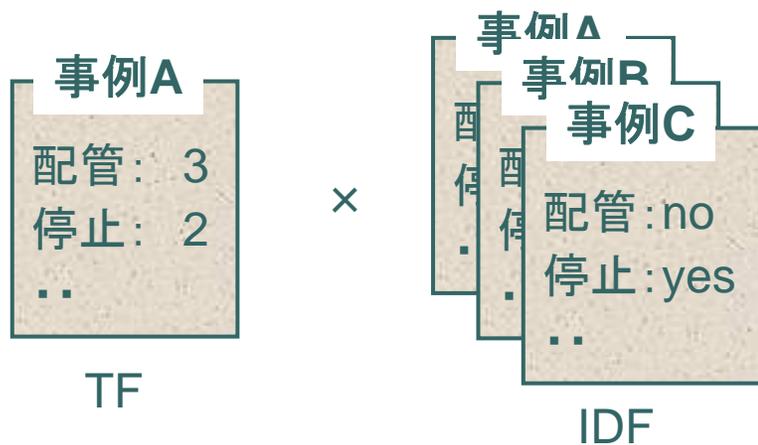
# 事例構造化

コンセプト別重要度 算出

コンセプト対関連度 算出

▶ **コンセプト別重要度** コンセプトの文書内頻度(TF)と総文書での使用度(IDF)の積

\*情報検索と言語処理 徳永健伸著  
言語と心理の統計 金明哲ほか



$$wei_k^d = tf(k, d) \cdot idf(k)$$

wei\_dk:

文書dにおけるコンセプトkの重要度

▶ **コンセプト対関連度**

- ・係り受け解析を利用した係り受け頻度
  - ・ハミング窓関数を利用したコンセプト出現密度
- } の積

# 事例構造化

コンセプト別重要度 算出

コンセプト対関連度 算出

## ▶ コンセプト対関連度

- ・係り受け解析を利用した係り受け頻度

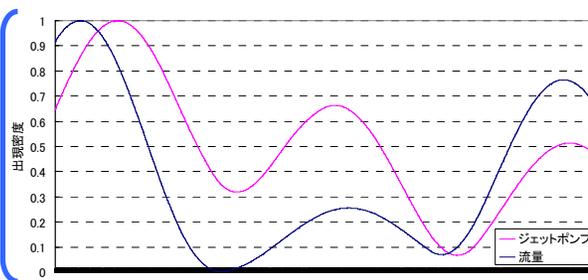
「配管」 → 「漏えい」: 4  
「原子炉」 → 「停止」: 1  
..

$$rel_{w1,w2} = freq\_pair(w1, w2)$$

- ・ハミング窓関数を利用したコンセプト出現密度

\*松倉健志 “文書の話題構造と文書間の意味的関連に基づくWeb検索に関する研究”(2001)

出現密度



コンセプトの出現密度をコサイン尺度によって類似度を算出

$$han_{w1,w2} = han\_weight(w1, w2)$$

指標の積 ( $rel_{w1,w2} * han_{w1,w2}$ ) をコンセプト対の関連度

### 事例A

点検を行っていたタービンからの蒸気を使って同タービンを駆動するタービンの小口径ドレン配管(同タービンを駆動した蒸気中に含まれる水分を回収して復水器へ移送するための配管)付近から漏えいしていることが確認された。このため、当該配管の詳細調査を実施することとし、3日午後10時50分から出力降下を開始し、4日午前7時48分に原子炉を手動停止した。なお、外部に及ぼす放射能の影響はない。調査の結果、4本ある当該ドレン配管のうち1本の配管が

# 事例構造化

## 事例A

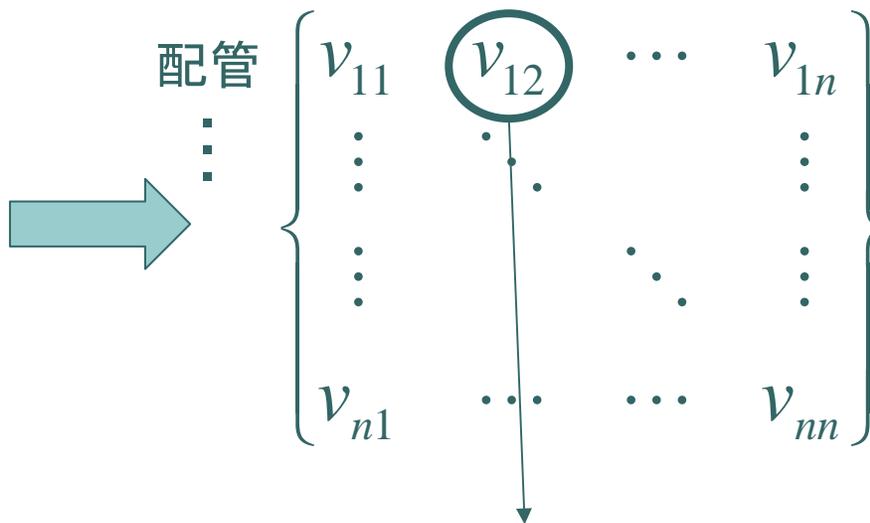
原子炉給水ポンプ駆動タービンドレン配管管台部からの漏えい  
 中部電力浜岡原子力発電所2号喫沸騰水型、定格出力84万ロワットは、定格出力にて運転中のところ、11月3日午後9時10分頃、タービン建屋3階にあるタービン駆動原子炉給水ポンプ室の床に水が溜まっていることが認められた。点検を行ったところ、高圧タービンからの蒸気を使って同ポンプを駆動するタービンの小口径ドレン配管(同タービンを駆動した蒸気中に含まれる水分を回収して復水器へ移送するための配管)付近から漏えいしていることが確認された。このため、当該配管の詳細調査を実施することとし、3日午後10時50分から出力降下を開始し、4日午前7時48分に原子炉を手動停止した。なお、外部に対する放射能の影響はない。中部電力浜岡原子力発電所2号喫沸騰水型、定格出力84万ロワットは、定格出力にて運転中のところ、11月3日午後9時10分頃、タービン建屋3階にあるタービン駆動原子炉給水ポンプ室の床に水が溜まっていることが認められた。点検を行ったところ、高圧タービンからの蒸気を使って同ポンプを駆動するタービンの小口径ドレン配管(同タービンを駆動した蒸気中に含まれる水分を回収して復水器へ移送するための配管)付近から漏えいしていることが確認された。このため、当該配管の詳細調査を実施することとし、3日午後10時50分から出力降下を開始し、4日午前7時48分に原子炉を手動停止した。なお、外部に対する放射能の影響はない。調査の結果、4本ある当該ドレン配管のうちの1本の管台部(配管が給水ポンプ駆動タービンに接続されている部分)に、直径約5mmの貫通孔が確認された。また、当該部及び他の小口径ドレン配管の管台部に、浸食と腐食の繰返しによる

点検  
 調査  
 1  
 対  
 正  
 11  
 階内  
 が  
 の  
 後  
 手  
 果、  
 を  
 た  
 上。  
 隙  
 時  
 そ

の

nは文書に現れる総単語数

.. 漏えい ...



Value<sub>1,2</sub>は配管と漏えいの総合関連度

$$Value_{1,2}^A = wei_1^A \cdot wei_2^A \cdot rel_{1,2}^A \cdot han_{1,2}^A$$

重要度

×

関連度

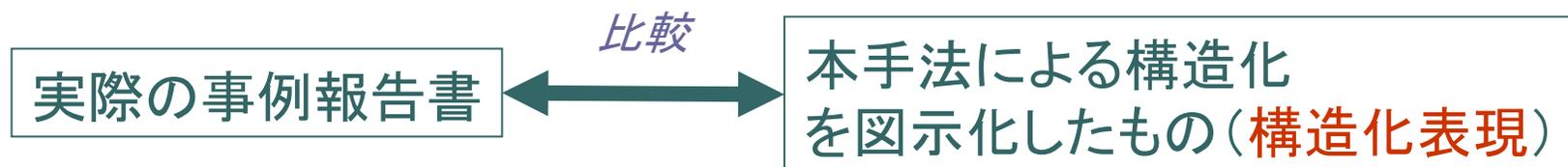
# ●●● | 事例構造化の評価

## ➡ 評価対象

1998年から2004年までで最も頻度の大きい「漏えい」に関する事例群  
27事例

「漏えい」キーワード周辺部の構造化を評価する

## ➡ 評価方法



### 構造化表現の評価項目:

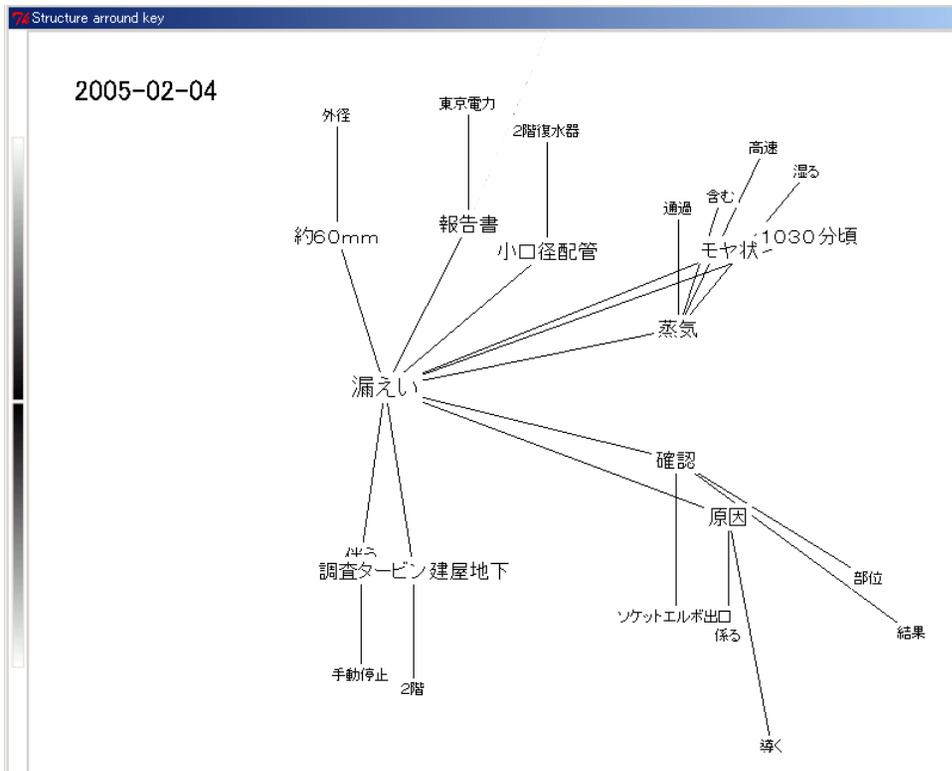
- ・「漏えい」がどこで起こったか(Where)
- ・何が「漏えい」したのか(What)

# 事例構造化の評価

例.2005.02.04 タービン建屋内小口径ドレン配管からの蒸気漏えい 柏崎刈羽1号機

## 構造化表現

## 事例報告書



原子力安全・保安院は、本日(平成17年2月4日)、東京電力から、柏崎刈羽原子力発電所1号機(沸騰水型:定格電気出力110万キロワット)のタービン建屋地下2階における蒸気漏れに伴う原子炉の手動停止について、以下のとおり報告を受けた。(東京電力からの報告内容) 柏崎刈羽原子力発電所1号機は定格熱出力一定運転中のあるところ。本日10時30分頃、タービン建屋地下2階復水器(A)近くの小口径配管からモヤ状に蒸気が漏れしていることを確認した。このため、本事象は直ちに運転に影響を及ぼすものではないが、本日11時45分に原子炉の停止操作を開始し、点検を行うこととした。なお、本事象による外部への放射性物質の影響はない。定格熱出力一定運転中の柏崎刈羽原子力発電所1号機(沸騰水型:定格電気出力110万キロワット)において、タービン建屋地下2階復水器(A)近くの小口径配管からモヤ状に蒸気が漏れしていることを確認したことから、原子炉を手動停止した事象(平成17年2月4日発表済み)に関し、東京電力は、本日(平成17年3月11日)、原子力安全・保安院(以下「当院」という。)に対し、原因と対策に係る報告書を提出した。

1. 原因と対策に係る東京電力の報告書の要点 (1) 調査結果① 漏えい箇所の調査タービン建屋地下2階の復水器(A)付近を詳細に確認した結果、復水器近くにある小口径のドレン配管(外径:約60mm)※1からの漏えいであり、当該ドレン配管のソケットエルボ※2出口近くの直管部に、直径1mm程度のピンホール(貫通孔)を2箇所確認した。なお、当該ドレン配管については、これまで減肉管理を実施していたが、ピンホールが確認されたソケットエルボ出口近くの直管部については、溶接箇所が近いため、超音波を用いた肉厚測定が難しい部位であった。また、ソケットエルボについても、超音波を用いた肉厚測定が難しい部位であった。② 当該ドレン配管の肉厚調査 当該ドレン配管におけるソケットエルボ及びその出口近くの直管部を中心に、超音波及び放射線を用いた肉厚調査を実施した結果、当該ドレン配管に設置されたオリフィス※3上流側においては、顕著な減肉傾向は確認されず、オリフィス下流側のソケットエルボとその出口近くの直管部においては、全般的に背側(曲り部の内面側)の減肉量が大きいことを確認した。③ 当該ドレン配管の内面観察調査 当該ドレン配管のオリフィス下流側を切断し、内面を詳細に調査した結果、復水器付近に設置されているソケットエルボ及びその出口近くの直管部の背側において、内部を通過する蒸気の流れによりエロージョン(浸食)が発生したものと考えられる模様を確認した。(2) 推定原因 当該ドレン配管に導かれた凝縮水を含む湿った蒸気が、オリフィスを通り、真空状態の復水器へ向かう過程で急激に減圧されるとともに、高速の蒸気流となった。この高速の凝縮水を含む蒸気流により、当該ドレン配管の内面にエロージョン(浸食)が発生し、徐々に配管の減肉が進展し、貫通に至ったため、蒸気が漏れ出したものと推定した。また、ソケットエルボ及びその出口近くの直管部は超音波を用いた肉厚測定が難しい部位であったため、直接、当該部位の点検を実施していなかったことも今回の事象が発生した一因だった。(3) 対策① 当該ドレン配管については、オリフィス下流から復水器

場所(Where):  
タービン建屋2階復水器の小口径配管  
ソケットエルボ付近の貫通孔

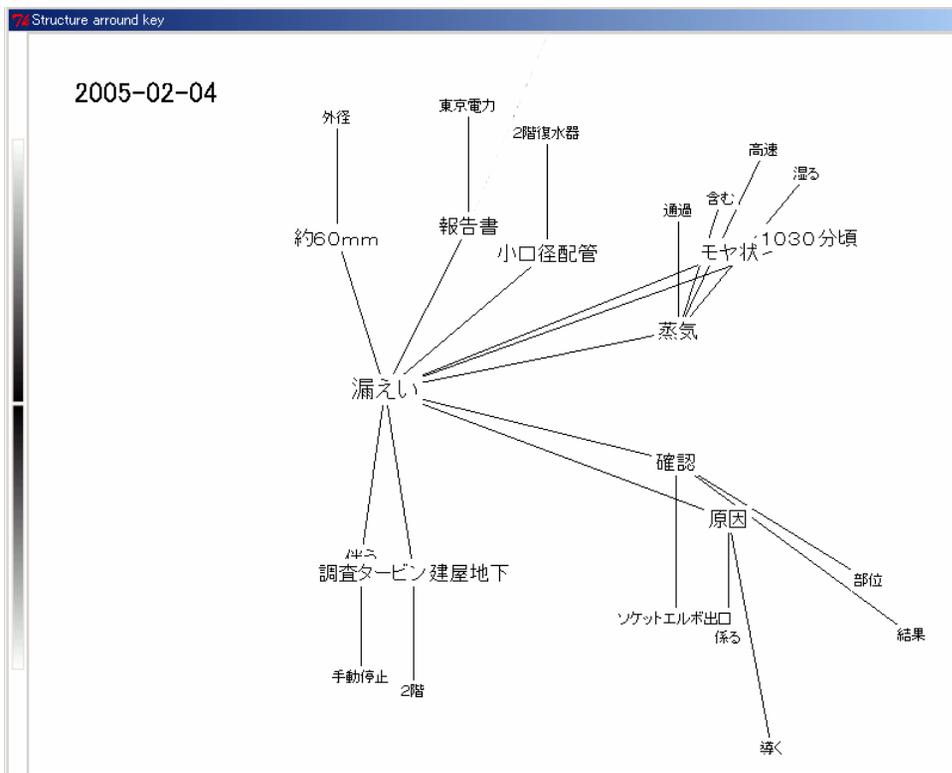
漏えい物質(What);  
モヤ状の蒸気

# 事例構造化の評価

例.2005.02.04 タービン建屋内小口径ドレン配管からの蒸気漏えい 柏崎刈羽1号機

## 構造化表現

## 評価結果



・「漏えい」がどこで起こったか  
(Where)

**完全に抽出**  
部分的に抽出  
不完全

・何が「漏えい」したのか  
(What)

**完全に抽出**  
部分的に抽出  
不完全

# ●●● | 事例構造化の評価

## ➡ 評価結果

評価項目 (What, Where) に関して構造化表現は {  
3: 完全に抽出  
2: 部分的に抽出  
1: 不完全

表.「漏えい」に関する事例群27事例の評価結果

	Where 現象発生機器の特定	What 何が漏えい
3の数(%)	77.78	62.96
2の数(%)	18.52	22.22
1の数(%)	3.70	14.81

部分的に抽出できた事例を含めると、Where 96%、What 85%となる。  
この結果より、「漏えい」の構造化は十分な精度をもっていると考えられる。

# ●●● | 事例知識の抽出

## ▶ 構造化した事例から事例知識を抽出

事例の特異性という観点から、事例知識を抽出しその有効性を評価する。

## ▶ 特異性について

特異な事例をただの一過性の事例としてとらえるべきでない

特異な事例を発見することによって、普遍的なトラブルでは気づかなかったプロセスの欠陥を見出し教訓化する。これによって、想定外のトラブル・事故を軽減することができる。

本研究では、特異な事例の発見に関わる事例特異性の数量的抽出を行う。

# ●●● | 事例特異性の導出

## ➡ 導出手順

$$Sim\_pair_A = \sum_{i=A}^Z \sum_{j,k=1}^n value_{j,k}^i$$

$pair(j,k)$ が1つ以上の事例で共有されるとき

$$Diff\_pair_A = \sum_{i=A}^Z \sum_{j,k=1}^n value_{j,k}^i$$

$pair(j,k)$ が上記以外するとき

$pair(j,k)$ は係り受けにある。係り*j*、受け*k*

事例A

$W_1 \rightarrow W_2$	:	$C_{12}$
$W_4 \rightarrow W_3$	:	$C_{43}$
$W_1 \rightarrow W_7$	:	$C_{17}$
$W_2 \rightarrow W_8$	:	$C_{28}$
$W_3 \rightarrow W_8$	:	$C_{38}$
$W_8 \rightarrow W_9$	:	$C_{89}$
$W_7 \rightarrow W_9$	:	$C_{79}$

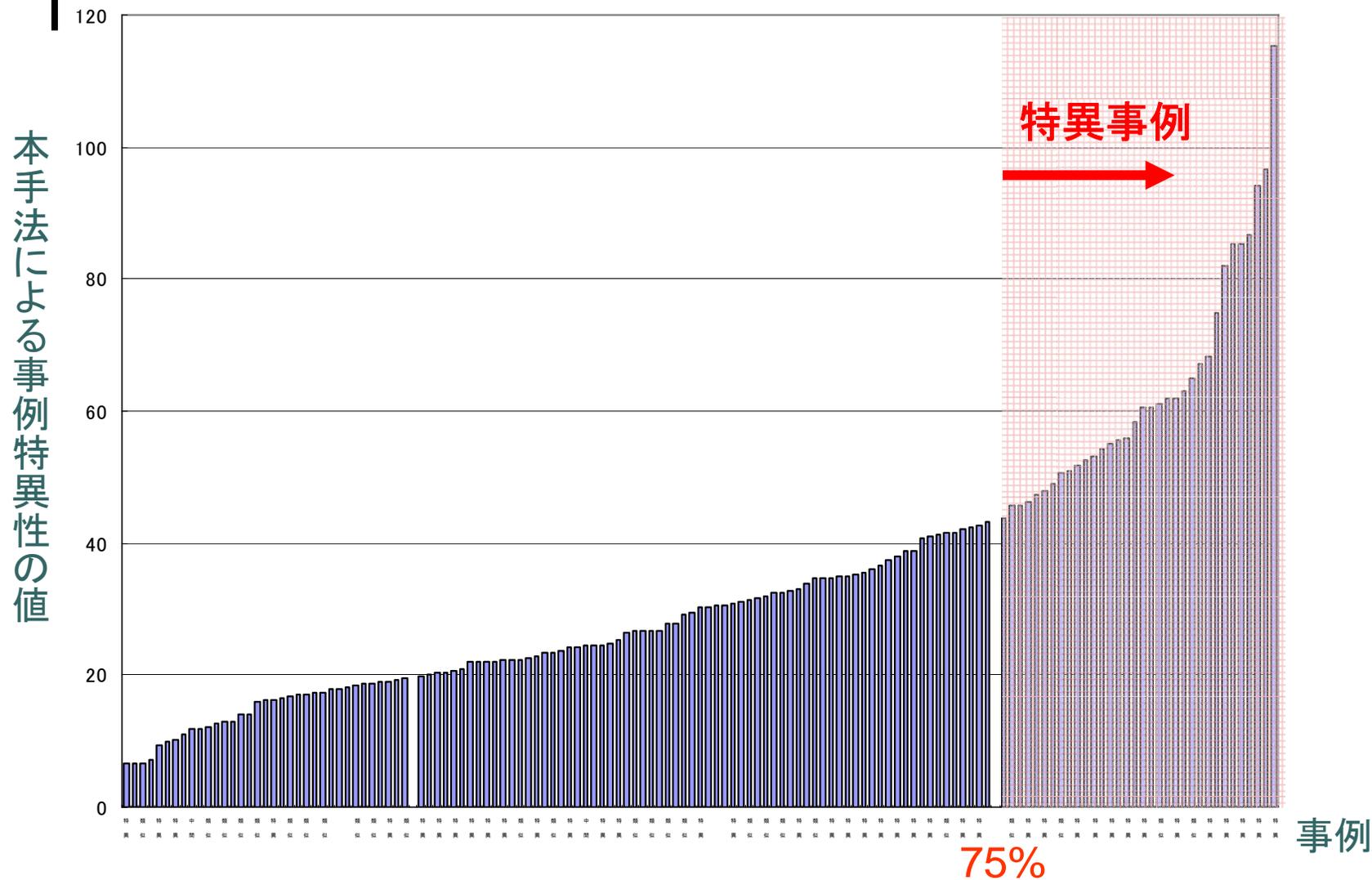
赤: 他事例で1つ以上共有されるペア

緑: それ以外的事例固有のペア

事例Aの特異性

$$Event\_diff_A = \frac{Diff\_pair_A}{Sim\_pair_A} \quad (\text{最大値で正規化})$$

# 本手法による事例特異性計算(昇順)



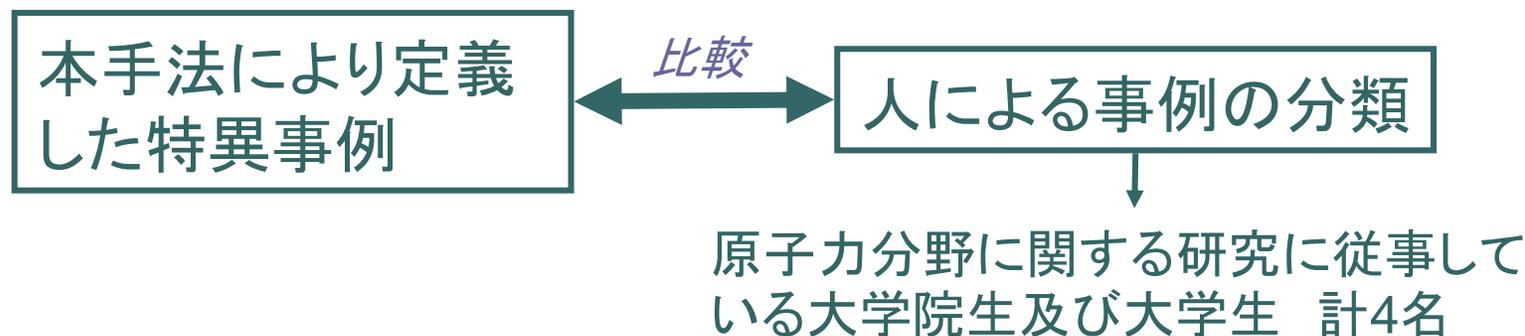
第3四分位数からの特異事例とした

# 事例特異性の評価

## 評価対象

1998年度から2004年度までの特異事例 34事例

## 評価方法



### ある事例が特異かどうかの判断手順

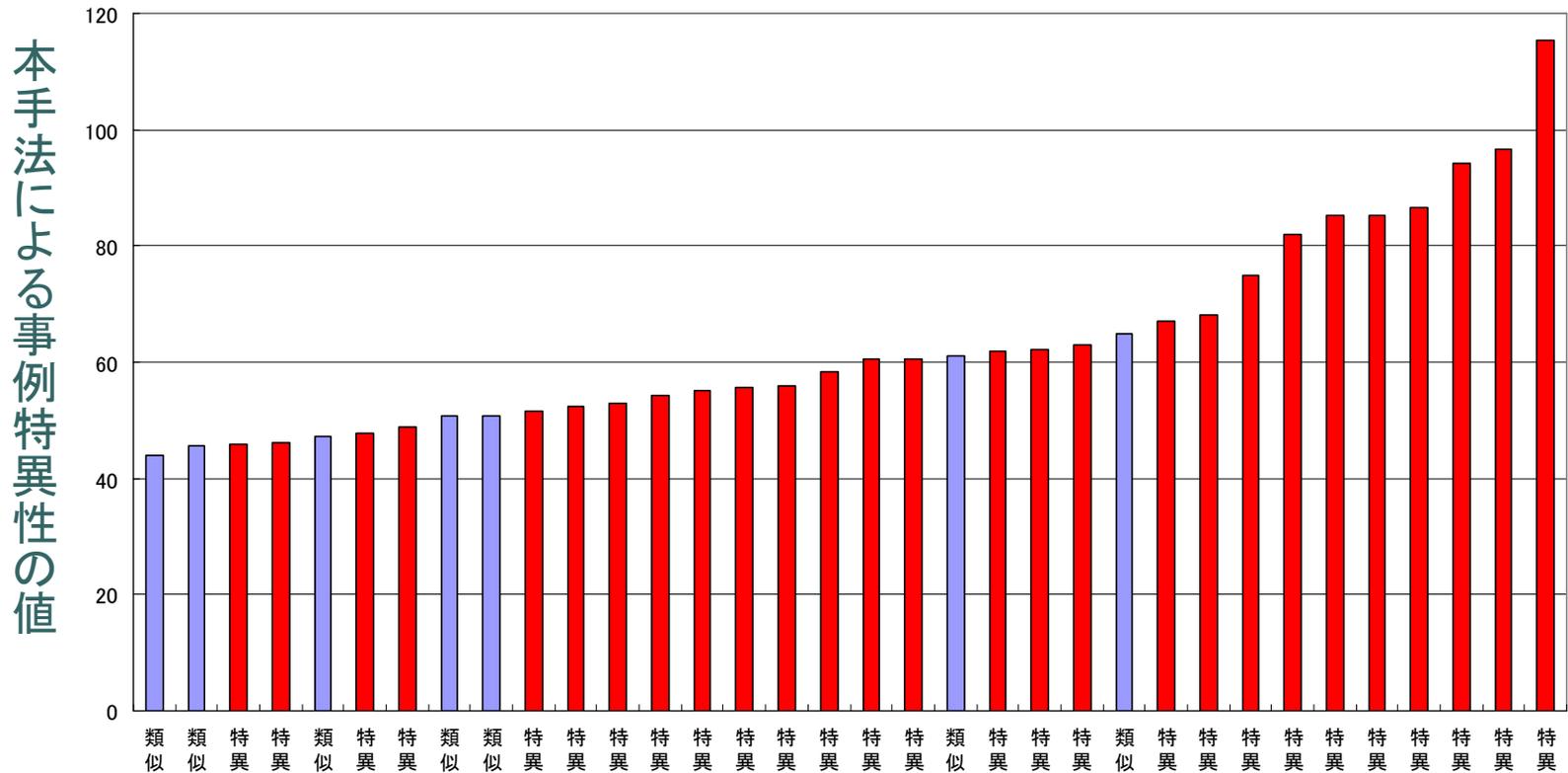
事例群の中で、他に類似している事例があれば類似カテゴリーに分類  
事例群の中で、それ以外(類似した事例がない)を特異カテゴリーに分類

判断材料: { 事例のタイトル  
概要

評価基準: { 現象  
事象発生機器

# 特異事例分類結果

人(MB)による特異な事例の分類結果



事例

赤で色分けされた事例は、人が特異であると判断した事例である  
青で色分けされた事例は、人が類似であると判断したものである

# ●●● | 事例特異性の評価

## ➡ 評価結果

表. 特異事例の中で人の判断と一致した割合(一致率)

協力者	一致率
MA	83.33
MB	78.13
BA	76.47
DA	63

4人の評価結果の平均をとった割合

**特異事例の中で人が特異だと判断した事例の割合: 75.3%**

特異事例に関する人の判断との一致率は8割弱となり、特異性算出の方法として十分な精度であると考えられる。

# ●●● | 結論

## ➡ 事例の解析可能な表現

本研究では、係り受け解析とハミング窓関数による単語出現密度を利用した多数の事例に適用が可能な事例構造化プロセスを提案した。さらに、「漏えい」に関する構造化の評価を行うことにより、本手法における事例構造化の十分な精度が示されたと考えている。

## ➡ 事例知識の抽出

事例特異性という観点から見たとき、本手法による事例知識抽出は十分な精度であると考ええる。

この十分性が示されたことによって、特異な事例の発見を行うことができる。その結果、特異性があるトラブルの未然防止につながると考えている。