

【公報種別】特許法第 17 条の 2 の規定による補正の掲載
 【部門区分】第 6 部門第 3 区分
 【発行日】平成 17 年 9 月 8 日 (2005.9.8)

【公開番号】特開 2002-245070 (P2002-245070A)
 【公開日】平成 14 年 8 月 30 日 (2002.8.30)
 【出願番号】特願 2001-42690 (P2001-42690)
 【国際特許分類第 7 版】

G 0 6 F 17/30

【 F I 】

G 0 6 F 17/30 2 2 0 Z

G 0 6 F 17/30 1 7 0 A

G 0 6 F 17/30 2 1 0 A

G 0 6 F 17/30 3 4 0 B

【手続補正書】

【提出日】平成 17 年 3 月 17 日 (2005.3.17)

【手続補正 1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

予め登録された文書の集合を対象として、その全体または一部分を取り出した処理対象文書集合に特徴的に出現する語句を抽出する特徴語句抽出ステップと、指定された要素を含む定義情報を設定する定義情報設定ステップと、該特徴語句抽出ステップで抽出した語句の中から該定義情報に含まれる各々の要素と関連する度合いが高い関連語句を取得する関連語句取得ステップを有するテキストマイニング方法において、

該定義情報に含まれる各々の要素間で関連語句を比較し、該定義情報に含まれる要素の多くと関連する関連語句を多重関連語句として抽出する多重関連語句抽出ステップを有することを特徴としたテキストマイニング方法。

【請求項 2】

請求項 1 記載のテキストマイニング方法において、

前記定義情報に含まれる各々の要素と関連する度合いが高い関連語句とは、該定義情報に含まれる各々の要素と予め定められた範囲内で共起する共起語句であることとし、

前記関連語句取得ステップは、前記特徴語句抽出ステップで抽出した語句の中から該定義情報に含まれる各々の要素と予め定められた範囲内で共起する共起語句を取得する共起語句取得ステップであり、

前記多重関連語句抽出ステップは、該定義情報に含まれる各々の要素間で共起語句を比較し、該定義情報に含まれる要素の多くと関わる共起語句を多重共起語句として抽出する多重共起語句抽出ステップであることを特徴としたテキストマイニング方法。

【請求項 3】

請求項 1 記載のテキストマイニング方法において、

前記処理対象文書集合は、語句または文章または文書を検索条件として用いて検索を行うことにより得られた文書集合であることを特徴としたテキストマイニング方法。

【請求項 4】

請求項 1 記載のテキストマイニング方法において、

前記多重関連語句抽出ステップは、前記定義情報に含まれる各々の要素間で関連語句を比較し、該定義情報に含まれる要素の多くと関連する関連語句を多重関連語句として取得

する多重関連語句取得ステップと、

該定義情報に含まれる各々の要素の関連語句から該多重関連語句を排除した要素関連語句を作成する多重関連語句排除ステップを有することを特徴としたテキストマイニング方法。

【請求項 5】

請求項 2 記載のテキストマイニング方法において、

前記多重共起語句抽出ステップは、前記定義情報に含まれる各々の要素間で共起語句を比較し、該定義情報に含まれる要素の多くと関わる共起語句を多重共起語句として取得する多重共起語句取得ステップと、

該定義情報に含まれる各々の要素の共起語句から該多重共起語句を排除した要素共起語句を作成する多重共起語句排除ステップを有することを特徴としたテキストマイニング方法。

【請求項 6】

請求項 5 記載のテキストマイニング方法において、

前記定義情報に含まれる要素の多くと関わる共起語句とは、予め定められた数以上の要素の共起語句として抽出された語句であることを特徴としたテキストマイニング方法。

【請求項 7】

請求項 5 記載のテキストマイニング方法において、

前記定義情報に含まれる要素の多くと関わる共起語句とは、該共起語句の該定義情報に含まれる各々の要素との共起の強さと共起する要素の数を基に算出した値が予め定められた値以上となる語句であることを特徴としたテキストマイニング方法。

【請求項 8】

請求項 6 記載のテキストマイニング方法において、

前記多重共起語句取得ステップは、予め定められた算出式で前記多重共起語句の重要度を算出する重要度算出ステップを有することを特徴としたテキストマイニング方法。

【請求項 9】

請求項 7 記載のテキストマイニング方法において、

前記多重共起語句取得ステップは、予め定められた算出式で前記多重共起語句の重要度を算出する重要度算出ステップを有することを特徴としたテキストマイニング方法。

【請求項 10】

請求項 8 記載のテキストマイニング方法において、

前記重要度は、前記多重共起語句が共起語句となった要素の数をを用いて予め定められた算出式で算出することを特徴としたテキストマイニング方法。

【請求項 11】

請求項 9 記載のテキストマイニング方法において、

前記重要度は、前記多重共起語句の前記定義情報に含まれる各々の要素との共起の強さと共起する要素の数をを用いて予め定められた算出式で算出することを特徴としたテキストマイニング方法。

【請求項 12】

請求項 4 記載のテキストマイニング方法において、

前記多重関連語句取得ステップで取得した多重関連語句と、前記多重関連語句排除ステップで作成した要素関連語句を提示する関連語句提示ステップを有することを特徴としたテキストマイニング方法。

【請求項 13】

請求項 5 記載のテキストマイニング方法において、

前記多重共起語句取得ステップで取得した多重共起語句と、前記多重共起語句排除ステップで作成した要素共起語句を提示する共起語句提示ステップを有することを特徴としたテキストマイニング方法。

【請求項 14】

請求項 8 または請求項 9 記載のテキストマイニング方法において、

前記多重共起語句取得ステップで取得した多重共起語句及び重要度と、前記多重共起語句排除ステップで作成した要素共起語句を提示する共起語句提示ステップを有することを特徴としたテキストマイニング方法。

【請求項 15】

予め登録された文書の集合を対象として、その全体または一部分を取り出した処理対象文書集合に特徴的に出現する語句を抽出する特徴語句抽出手段と、指定された要素を含む定義情報を設定する定義情報設定手段と、該特徴語句抽出手段で抽出した語句の中から該定義情報に含まれる各々の要素と関連する度合いが高い関連語句を取得する関連語句取得手段を備えるテキストマイニング装置において、

該定義情報に含まれる各々の要素間で関連語句を比較し、該定義情報に含まれる要素の多くと関連する関連語句を多重関連語句として抽出する多重関連語句抽出手段を備えることを特徴としたテキストマイニング装置。

【請求項 16】

予め登録された文書の集合を対象として、その全体または一部分を取り出した処理対象文書集合に特徴的に出現する語句を抽出する特徴語句抽出モジュールと、指定された要素を含む定義情報を設定する定義情報設定モジュールと、該特徴語句抽出モジュールで抽出した語句の中から該定義情報に含まれる各々の要素と関連する度合いが高い関連語句を取得する関連語句取得モジュールを含むテキストマイニングシステムを構築するためのプログラムを格納した記憶媒体において、

該定義情報に含まれる各々の要素間で関連語句を比較し、該定義情報に含まれる要素の多くと関連する関連語句を多重関連語句として抽出する多重関連語句抽出モジュールを含むことを特徴とした記憶媒体。