

テキストマイニングツール「ユーザーローカル AI テキストマイニング」を使う

KH Coder が有料ツールになったため、今後は「ユーザーローカル AI テキストマイニング」を使って分析する。

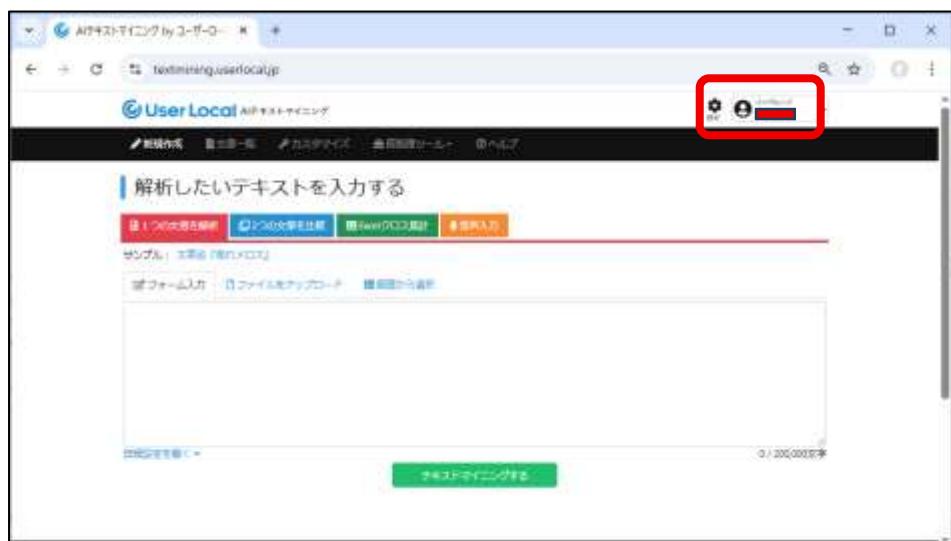
今回は、「さらば黄昏」(長塚圭史著)のテキストデータを題材にして、分析する。

1. User Local AI テキストマイニングに、用意したファイルを読み込ませる。

戯曲の全文章を打ち込んだテキストファイルを User Local AI テキストマイニングに読み込ませる。

User Local(<https://textmining.userlocal.jp/>)

自分のアカウントを作って、ログインしておくと大量の文字が扱えるようになる。



ファイルをドラッグする



詳細設定を開くを押す



辞書の作成・編集はこちら を押す



地名・人名などが複数の単語として分かれないように登録する

※そのまま分析すると踊田は踊と田に分かれる

カスタム辞書を作成する

固有名詞や除外語を設定して、テキストマイニングをカスタマイズできます。

基本設定

辞書名

さらば黄昏

単語設定

固有名詞設定

設定した単語を1つの言葉として認識させます。これにより「株式会社ユーザーローカル」などの固有名詞が分割されずに解析できます。なお、各単語は改行区切りで入力してください。

踊田
竹井大
本町一栄
阪本慎平
阪本雪乃
阪本朝陽
剣持愛一郎
大河原
島村潤子
自走

14 / 30語

設定するボタンをクリックする

除外語設定

設定した単語を解析結果から除外します（=ストップワード）。これは、出現頻度が高い割には役に立たない単語（「ない」「思う」など）を除くために行います。なお、各単語は改行区切りで入力してください。

単語1
単語2
単語3

0 / 30語

同義語設定

解析時に同じ意味を持つ単語を1つにまとめます（例：同義語「パソコン」「PC」→代表語「パーソナルコンピューター」）。

同義語
同義語
+ 同義語を追加

代表語

+ 同義語設定を追加

設定する

同義語: 0 / 30語

解析したいテキストを入力する

サンプル：太宰治『走れメロス』

フォーム入力 ファイルをアップロード 履歴から選択

以下の形式がアップロードできます。最大200,000文字、10MBまで。
テキスト形式(.txt)、CSV形式(.csv)、Word形式(.docx)、Excel形式(.xlsx .xlsm)、PowerPoint Text形式(.pptx)、PDF形式(.pdf)

解析対象に名前を設定（任意）
「さらば黄昏」全6場（分析用）.txt

解析するファイル



辞書設定（辞書の作成・編集はこちら）
さらば黄昏

公開設定 ?
 外部に公開 URLを知っている全員が閲覧可能
 公開しない 自分のみ閲覧可能

詳細設定を開じる ^

テキストマイニングする

テキストマイニングするボタンを押すと結果が表示される

ワードクラウド

スコアが高い単語を複数選び出し、その値に応じた大きさで図示しています。単語の色は品詞の種類で異なっており、青色が名詞、赤色が動詞、緑色が形容詞、灰色が感動詞を表しています。[【このテキストマイニング結果を利用するには】](#)



※以降、分析結果は

※ユーザー一ローカル AI テキストマイニングによる分析(<https://textmining.userlocal.jp/>)

単語出現頻度

文章中に出現する単語の頻出度を表にしています。単語ごとに表示されている「スコア」の大きさは、与えられた文書の中でその単語がどれだけ特徴的であるかを表しています。通常はその単語の出現回数が多いほどスコアが高くなりますが、「言う」や「思う」など、どの文書にもよく現れる単語についてはスコアが低めになります。【スコアの目的と算出方法について】

名詞	スコア	出現頻度
竹井	5979.44	468
剣持	2141.86	241
輝	707.12	136
慎平	1954.99	133
雪乃	1085.93	131
大河原	1264.73	126
潤子	1340.14	118
朝陽	898.66	114
一栄	1618.54	112
躊躇	828.06	61
住民	218.22	56
男	19.51	52
犬塚	441.57	50
阪本さん	299.02	30
電話	11.13	30

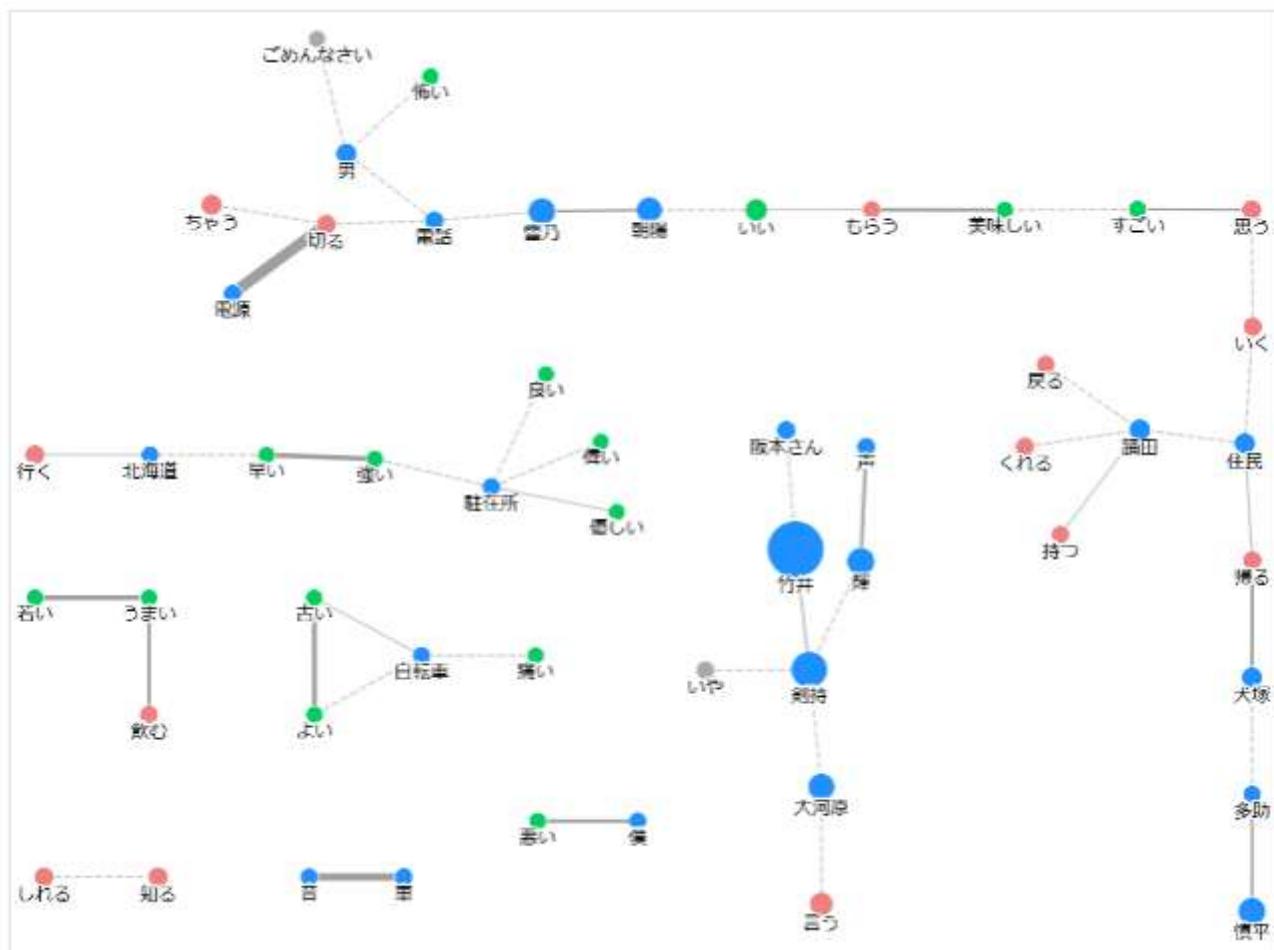
動詞	スコア	出現頻度
言う	3.94	69
わかる	6.54	59
ちゃう	3.28	43
知る	2.87	34
行く	0.83	33
思う	0.58	32
しれる	5.18	31
切る	10.13	30
考える	2.20	28
いく	1.40	27
帰る	1.60	25
聞く	1.40	24
くださる	1.05	23
くれる	0.63	23
戻る	3.77	21

形容詞	スコア	出現頻度
いい	2.27	55
すごい	0.38	13
美味しい	0.36	9
悪い	0.30	9
怖い	0.25	7
よい	0.10	7
若い	0.75	6
痛い	0.12	5
強い	0.11	5
早い	0.07	5
うまい	0.16	4
良い	0.02	4
古い	0.43	3
赤い	0.38	3
偉い	0.33	3

感動詞	スコア	出現頻度
いや	6.70	26
ごめんなさい	1.20	9
ありがとう	0.09	8
すいません	1.80	7
おい	0.74	5
じゃあ	0.90	2
いいえ	0.72	2
うお	0.36	2
そうですね	0.31	2
なるほど	0.05	2
お疲れ様	0.02	2
ごめんください	4.14	1
はいはい	0.40	1
さらば	0.29	1
いえいえ	0.12	1

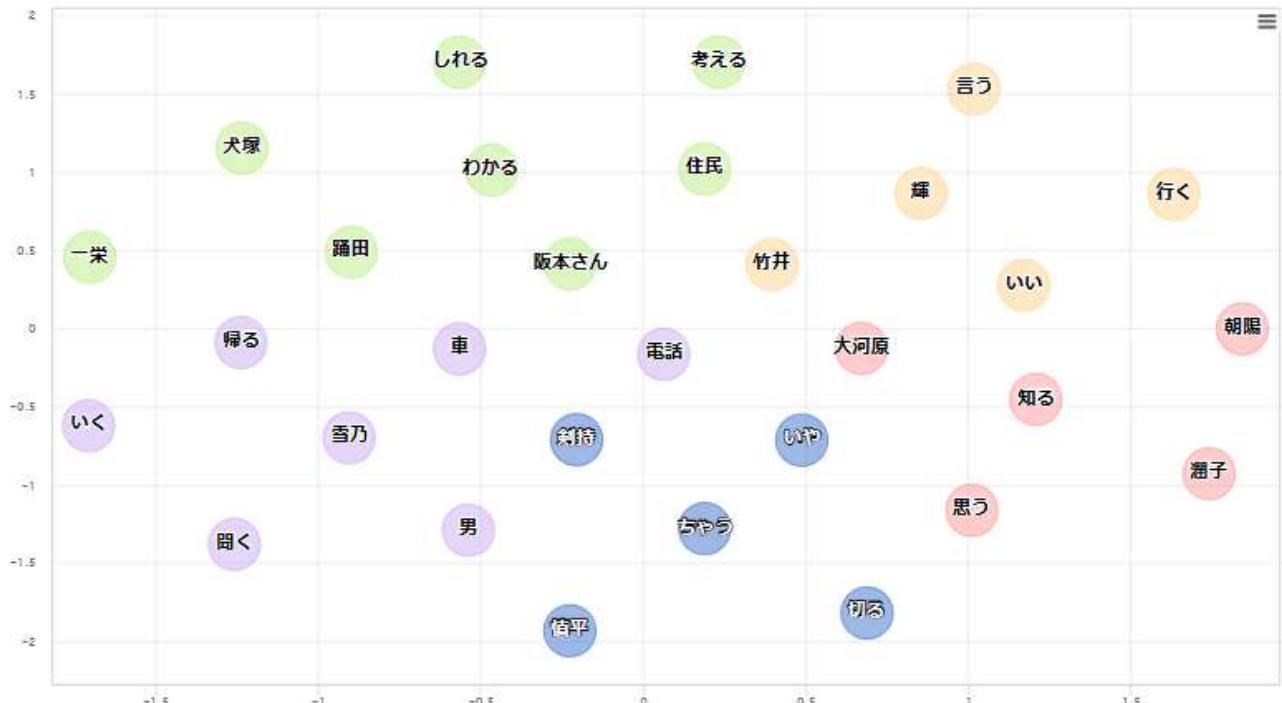
共起キーワード

文章中に出現する単語の出現パターンが似たものを線で結んだ図です。出現数が多い語ほど大きく、また共起の程度は強い方から順に 太い実線 > 細い実線 > 破線 で描画されます。[【共起とは?】](#)



2次元マップ

文章中での出現傾向が似た単語ほど近く、似ていない単語ほど遠く配置されています。距離が近い単語はグループにまとめ、色分けしています。[【2次元マップの詳しい説明】](#)



係り受け解析

係り受け解析では、「名詞」に係る「形容詞」「動詞」「名詞」についての解析結果を表示します。「スコア」は、出現回数やその係り受け関係が全組み合わせのうちに占める割合などを複合的に判断し、独自に算出した数値です。「スコア」が高いほど、よりその係り受け関係が重要であることを示します。また、単語の後に「(否: 50%)」などとついている場合、集計された係り受け関係のうち 50%が否定表現(例:「高い」→「高くない」)として使われていることを意味しています。ネガポジは名詞にかかる形容詞がポジティブ(ネガティブ)な単語かどうかを表しています。[【係り受け解析とは?】](#)
(結果の表示に時間がかかる場合があります)

■ 名詞 - ■ 形容詞

名詞 - 形容詞	スコア	出現頻度
竹井 - いい	0.36	4
輝 - いい	0.21	3
雪乃 - 怖い	0.75	2
僕 - 悪い	0.60	2
一栄 - すごい	0.43	2
大河原 - いい	0.11	2
パトロール - いい	0.11	2
男 - おぞましい	1.00	1
頭 - ずるい	1.00	1
剣持 - 違い	1.00	1
ママ - 悲しい	1.00	1
輝 - 近い	1.00	1
道内 - あったかい	中立	1
雪 - 少ない	中立	1
一栄 - 可愛い	ポジティブ	1

■ 名詞 - ■ 動詞

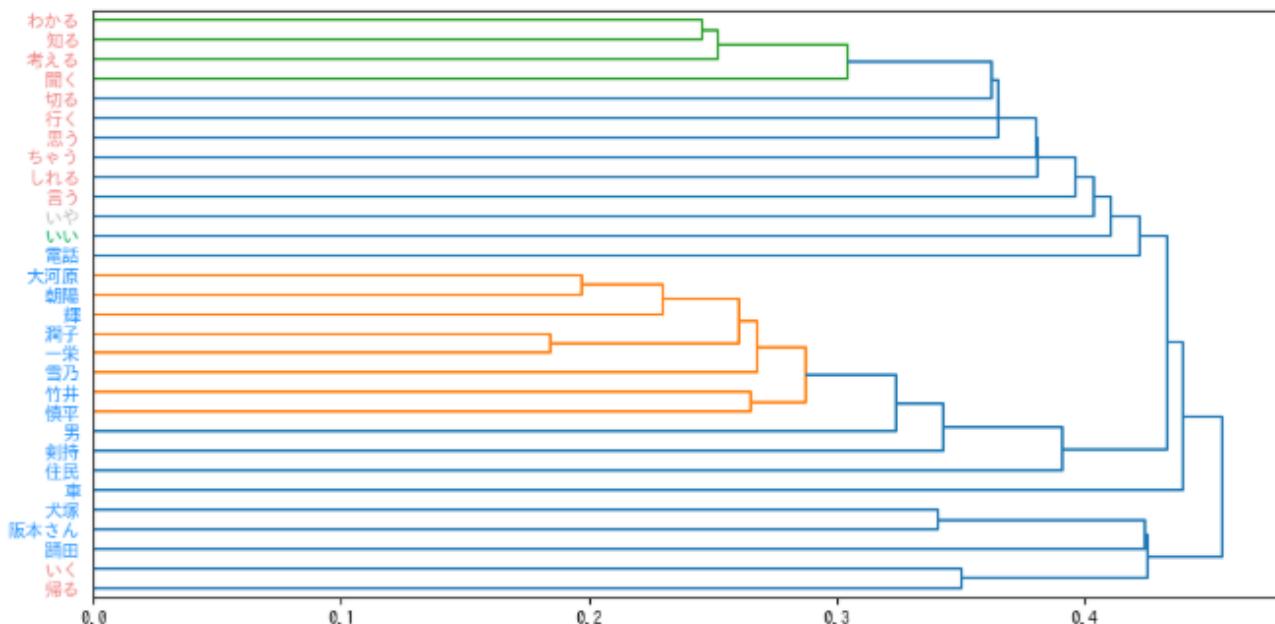
名詞 - 動詞	スコア	出現頻度
毛源 - 切る (否: 22.22%)	2.90	9 (否: 2)
竹井 - わかる (否: 42.86%)	0.93	7 (否: 3)
剣持 - 答える	4.20	6
質問 - 答える	4.20	6
竹井 - 行く (否: 16.67%)	1.24	6 (否: 1)
大河原 - 言う	0.60	6
竹井 - 言う (否: 16.67%)	0.60	6 (否: 1)
剣持 - くださる	1.25	5
竹井 - 見る	1.03	5
竹井 - しれる (否: 100.00%)	0.94	5 (否: 5)
剣持 - 行く	0.88	5
輝 - 来る (否: 60.00%)	0.70	5 (否: 3)
竹井 - 出る (否: 20.00%)	0.56	5 (否: 1)
潤子 - 出る	0.56	5
声 - かける	1.67	4

■名詞 - ■名詞

名詞 - 名詞	スコア	出現頻度
剣持 - 大丈夫	2.55	7
雪乃 - 朝陽	0.26	5
剣持 - 輝	0.22	5
竹井 - 北海道	1.33	4
輝 - 大丈夫	0.91	4
竹井 - 剣持	0.08	4
剣持 - 仕事	1.20	3
踊田 - 全員	1.00	3
雪乃 - 連絡	1.00	3
竹井 - 連絡	1.00	3
車 - 音	0.75	3
一栄 - 大丈夫	0.55	3
輝 - 声	0.52	3
雪乃 - 電話	0.39	3
竹井 - 電話	0.39	3

階層的クラスタリング

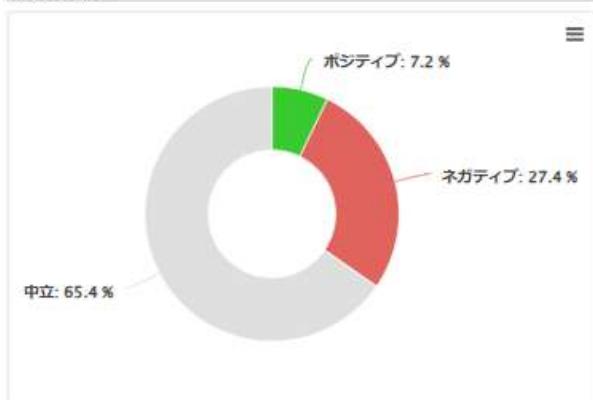
文章中での出現傾向が似た単語をまとまりとしてとらえられるよう樹形図で表したものです。グループは色分けして表示しています。[【階層的クラスタリングの詳しい説明】](#)



サマリー

文書全体を分析し、感情の傾向を可視化しています。「ポジネガ」は、文章に含まれるポジティブな感情の文とネガティブな感情の文の存在比を示しています。「感情」は、文章に含まれる各感情の度合いを数値に換算しています。なお、各感情の数値は、全ての感情の平均値を 50%とした偏差値です。

ポジネガ



ポジネガ推移

文書を分割して分析し、ポジティブな感情とネガティブな感情の比率の推移を時系列的に可視化しています。グラフの左端が文書の先頭を表しており、右に進むほど終わりに近づきます。



AI 文章要約

ChatGPT のアルゴリズムと連携し、高精度な要約を生成します。15~60 秒かかることがあります。

※入力されたデータが、AI 学習に利用されることはありません

この作品「さらば黄昏」は、登場人物たちの人間関係や地域社会の問題を描いたドラマです。物語は、ひき逃げ事件をきっかけに、地域住民が集まり、事件の真相やその影響について議論を交わす様子から始まります。1. **事件の発端**: 竹井大という駐在員が、ひき逃げに遭った阪本朝陽とその母親雪乃から話を聞く場面が描かれ、朝陽が自転車で避けた際に車に巻き込まれたことが語られます。「朝陽はライトつけてますよ。つけてたでしょ?」というやり取りから、事件の詳細が徐々に明らかになります。2. **地域の反応**: 事件を受けて、地域住民が集まり、竹井がパトロール隊の結成を提案します。「この踊田地区を、このまま過ごしやすい穏やかな集落にしたいじゃないですか」と、地域の安全を守るための行動が求められます。3. **住民の対立**: 住民たちの間には意見の対立が生じ、「暴力反対!」と叫ぶ声が上がる中で、竹井は冷静に地域の分断を防ぐための提案を続けます。「このまま何もせず事態を放置しておけば、やがてその時が来た時、ビクビク怯えて…」と、地域の未来を憂慮する姿が描かれます。4. **個人の葛藤**: 朝陽は自らの責任を感じ、「僕が悪いんです」と発言し、事件の影響を受けたことを示します。母親の雪乃是感情的になりながらも、息子を守ろうとする姿が印象的です。5. **結末と未来への希望**: 最終的に、朝陽や他の住民たちがパトロール隊に参加する意志を示し、地域の安全を守るための行動を起こすことが決まります。「僕／私やります」との言葉が、未来への希望を象徴しています。物語は、個々の葛藤や地域の問題を通じて、コミュニティの絆や協力の重要性を描いています。

(以上、User Local によるテキストマイニング結果)

ChatGPT よる要約が的外れだったので、この文章と戯曲全文を Gemini に読ませて、概要を批判せよ。

■Gemini による要約

ご提示いただいた要約文を読むと、確かに「的を外している」と感じられるのも無理はありません。その要約は、物語の表面的な言葉(パトロール、安全、希望)だけを拾ってしまい、この戯曲の核心である**「過去の因縁」「暴力の影」「人々の無責任さと欺瞞」といったドロドロとした緊張感**が完全に抜け落ちてしまっています。

特に「未来への希望を持たせる形で締めくくられる」という部分は、あの静かで不穏なラストシーンを読み解けていない印象です。

私(Gemini)が、全六場を通読した上での「物語の本質」を突いた要約を作成します。

戯曲『さらば黄昏』要約

【概要】 本作は、のどかな田舎町「踊田地区」を舞台に、25 年前の凄惨な放火殺人事件の記憶と、その加害者の帰還という影に怯える人々の姿を描いた心理サスペンスである。正義を信じようとする駐在員と、過去に縛られた住民たちの対比を通じ、平穏なコミュニティの裏側に潜む排他性と無責任さを浮き彫りにする。

1. 繰り返される暴力の予兆 物語は、少年・朝陽が「正体不明の車」に煽られ負傷する事件から幕を開ける。駐在員の竹井は、これがかつて自分が逮捕した凶悪犯・犬塚の親族による「警告」ではないかと直感する。竹井は定年退職と同時に被害者の娘・一栄と北海道へ移住する計画を立てていたが、村に忍び寄る不穏な空気が彼を足止めする。

2. 英雄の孤独と住民の無関心 竹井は村を守るために「自警団(パトロール隊)」の結成を住民に促すが、住民たちは竹井を「かつて発砲した危ない男」と揶揄し、自分たちがリスクを負うことを拒絶する。守ろうとする竹井の「正義」は空回りし、住民たちの「人任せ」な態度との溝は深まっていく。

3. 暴かれる「血」の秘密 竹井の後任予定である剣持と、犬塚の元愛人・潤子の娘である輝の結婚話が進む中、衝撃的な事実が明かされる。輝は「犬塚の娘」である可能性が高く、その血縁関係が村のパワーバランスを揺さぶり始める。深夜、駐在所に現れた謎の男は、圧倒的な「悪」の気配で村の平和が砂上の楼閣であることを突きつける。

4. 崩壊する仮面 パトロール隊の結成集会で、住民たちの不満が爆発する。竹井を「セイロン(異分子)」と呼び排除しようとする醜い集団心理が露呈する中、良心的な市民であったはずの慎平が、自身の過去の罪や輝との繋がりを暗示して逃亡するように姿を消す。竹井は、自分が守ろうとした「平和な村」など最初から存在せず、全員が共犯者として秘密を抱えていたことに気づかされる。

5. 結末:沈黙と停滞 数年後。竹井と一栄は北海道へ去り、大河原など古い世代は世を去った。村には静かな日常が戻ったように見えるが、それは「解決」ではなく、単なる「忘却と沈黙」による安定であった。残された慎平と一栄が交わす短い言葉には、癒えることのない傷跡と、一生消えない「罪」の感覚が重く横たわっている。

Gemini の分析:なぜ前の要約はダメだったのか?

前の要約が「的外れ」だった理由は以下の 3 点に集約されます。

1. 「犬塚」という不在の主役を無視した: この劇は「犬塚が帰ってくる」という恐怖がすべての動機になっています。それを見落とすと、単なる「防犯意識向上の物語」に見えてしまいます。
2. ジャンルの誤解: これを「社会派の感動ドラマ」として処理したため、「希望」という言葉を使ってしまった。実際には「ウェスタン風の心理スリラー」であり、結末は「ビター(苦い)」なものです。
3. 慎平の行動の意味を捉えていない: 慎平がトイレに行くと嘘について逃げたのは、感動的な場面ではなく、この村の「闇」が限界に達した瞬間です。

あなたがタイピング中に感じた**「住民の対比が面白い」「竹井も問題がある」「慎平が不可解」**という感覚こそが、この戯曲の正解だと思います。私の要約では、その「不自然さ」を強調するように構成しました。

(以上、Gemini による要約)

Gemini はただテキストを読ませただけでなく、事前に戯曲について会話をしているため、その情報を汲んで訳している Gemini の方が、よりよい要約になっている。ただし、大河原家の葬式は、おそらく村の最年長の大河原家の姫子さんのお葬式だと思うので、誰が死んだか把握できていない。あと、朝陽のセリフは、男の場合のセリフ／女の場合のセリフとして書いて置き、共通するセリフをそのままつなげていて、機械的に判別しにくい。もしかすると、AI はそれが読み取れず、朝陽を男と認識して回答している可能性ある。

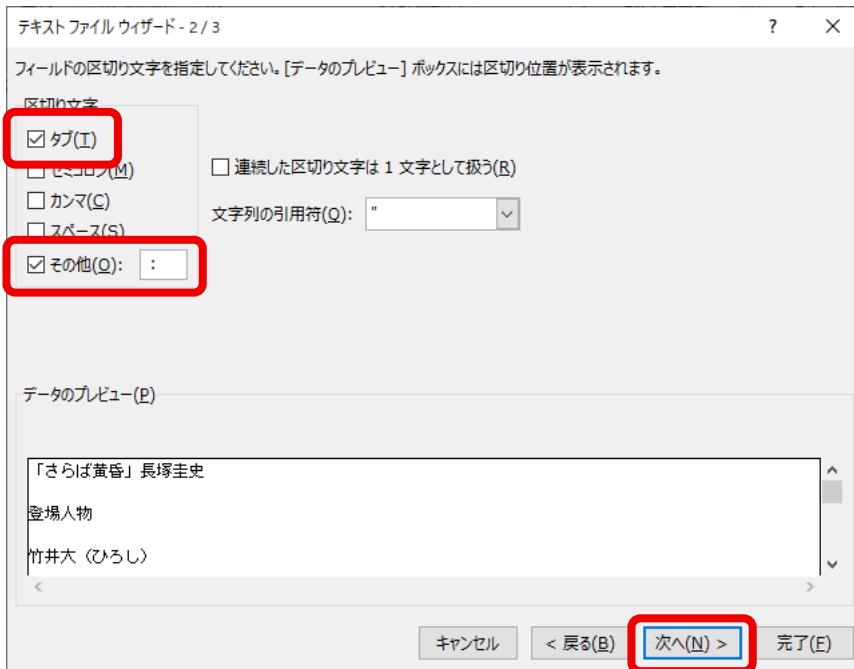
2. テキストデータを Excel で整形する。

ある人物のセリフのみにしたいというときなどは、Excel などの表のデータにしておけば、フィルタ機能を使って、それぞれのセリフを容易に取り出せる。

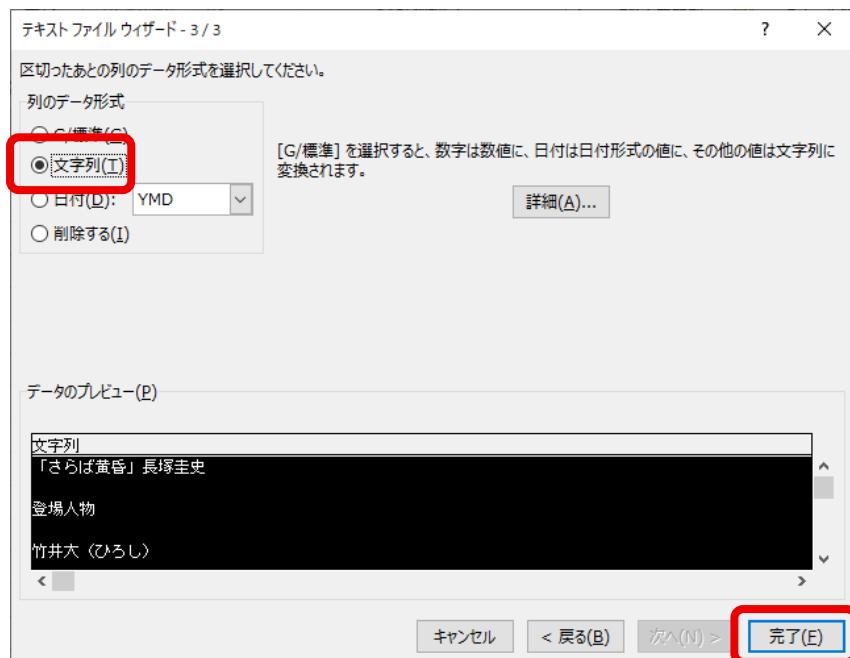
Excel で役名とセリフに分けやすくするために、役名とセリフの間の全角スペースを「:」(全角コロン)に置換する。メモ帳の置換機能を使い、「役名」を「役名:」とするように置換すればいい。置換後は漏れがないか確認する。

スペース区切り	→	コロン区切り
竹井 無灯火だった? 雪乃 朝陽はライトつけて 朝陽 つけてた。 雪乃 え？ つけてたよね 朝陽 つけてたつけてた。	→	竹井:無灯火だった? 雪乃:朝陽はライトつけて 朝陽:つけてた。 雪乃:え？ つけてたよね 朝陽:つけてたつけてた。

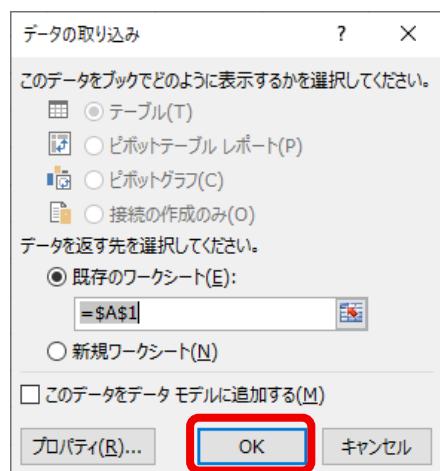
新しく Excel ファイルを作り開き、「データ」メニューの「外部データの取り込み」から、戯曲のテキストファイルを選び、区切り文字を「タブ」とその他「:」にする。ト書きの前にはタブを入れていたのでト書きは B 列に表示される。役名は A 列に、セリフは B 列に来る。文章中に「:」があった場合、区切られるので、整形後、Excel ファイルを一通りチェックする。



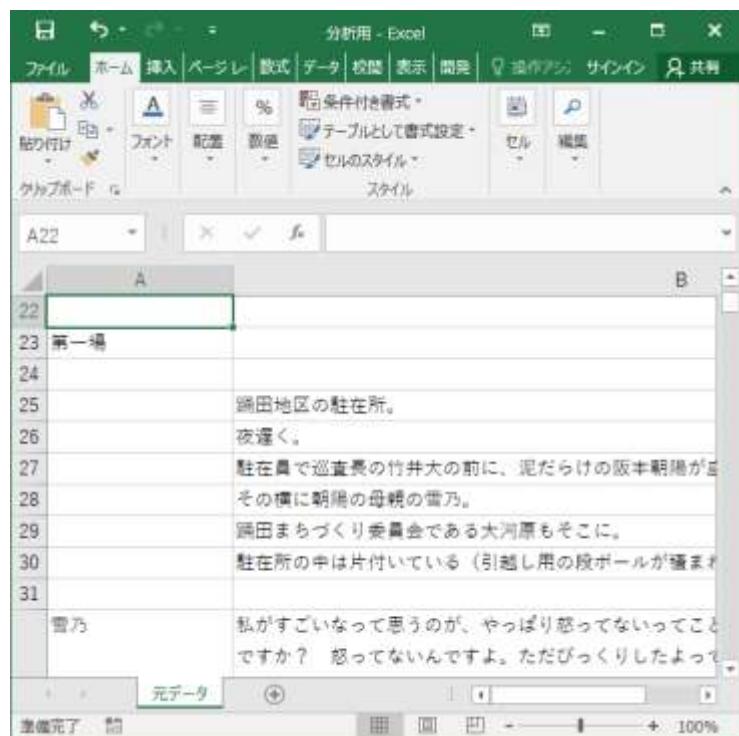
データ形式は念のため文字列にした。今回の戯曲なら標準にしても取り込んだ後、書式がおかしくなる可能性は少ない。「完了」ボタンを押す。



そのままそのブックの A1 に結果を入れればいいので、OK ボタンをクリックする。



Excel ファイルに戯曲のテキストが読み込まれる



The screenshot shows an Excel spreadsheet titled '分析用 - Excel'. The data is organized into two columns: Column A (役名) and Column B (セリフ). The first few rows are as follows:

役名	セリフ
第一場	
	踊田地区の駐在所。
	夜遅く。
	駐在員で巡査長の竹井大の前に、泥だらけの阪本朝陽が立つ。
	その横に朝陽の母親の雪乃。
	踊田まちづくり委員会である大河原もそこに。
	駐在所の中は片付いている（引越し用の段ボールが積まれている）。
雪乃	私がすごいなって思うのが、やっぱり怒ってないってことですか？ 怒ってないんですよ。ただびっくりしたよって

これで、B 列からセリフの前の役名がないデータを取り出すことができ、頻出語も役名が回数が多くなったが、それを回避することができる。

セリフのみのデータの分析結果

ワードクラウド



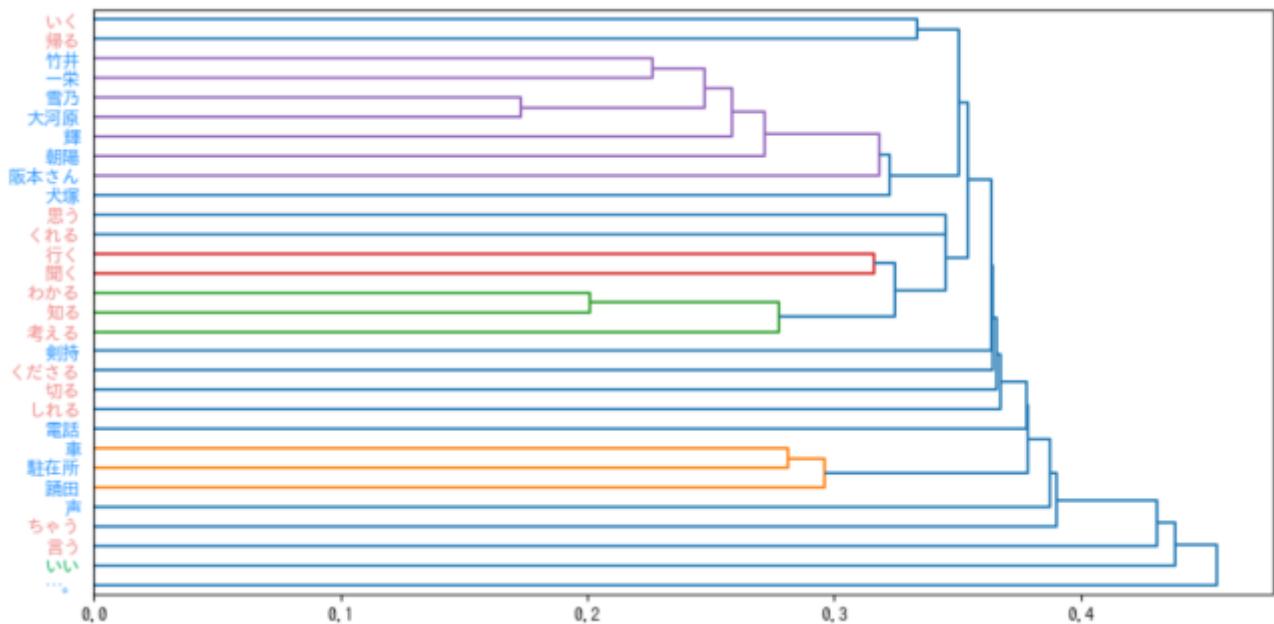
単語出現頻度



共起キーワード

A	B	C	D
単語ペア	単語1	単語2	共起回数
1 切る 電源	切る	電源	17
2 地区 踏田	地区	踏田	9
3 刺持 繩	刺持	縄	9
4 刺持 竹井	刺持	竹井	8
5 北海道 竹井	北海道	竹井	7
6

階層的クラスタリング



3. 特定の人物のセリフを分析する

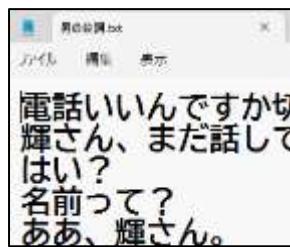
2列の表にしてフィルタ機能を使って、A列のフィルタボタンで登場人物を絞り込めるようにする。

A	B
1 場・登場人物 ▾	ト書き・台詞
2 第一場	
3	
4	踊田地区の駐在所。
5	夜遅く。
6	駐在員で巡査長の竹井大の前に、泥だらけの阪本朝陽が座っている。
7	その横に朝陽の母親の雪乃。
8	踊田まちづくり委員会である大河原もそこに。
9	駐在所の中は片付いている（引越し用の段ボールが積まれていてもよ
10	
11 雪乃	私がすごいなって思うのが、やっぱり怒ってないってことだと思うん 怒ってないんですよ。ただびっくりしたよって。でも自転車のタイヤ す。ぐにゃんぐにゃんってなっちゃって。だからわあって避（よ）け でも普通あれですよね、自転車が避けるわけですからこっち傾くわけ んですから、向こうは一切ハンドル切ってない、もしくは若干朝日側 それ避ける朝陽の動体視力すごくないかって思っちゃってるんですけ
12 竹井	無灯火だった？
13 雪乃	朝陽はライトつけてますよ。つけてたでしょ？

男のセリフだけ取り出したければ、1行目のボタンを押して、男だけチェックする。

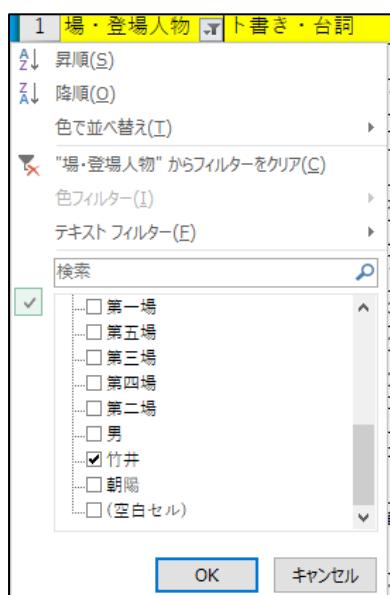
A	B
1 場・登場人物 ▾ ト書き・台詞	
昇順(S)	
降順(O)	
色で並べ替え(I)	
"場・登場人物" からフィルターをクリア(C)	
色フィルター(I)	
テキスト フィルター(E)	
検索 <input type="text"/>	
<input type="checkbox"/> 第一場 <input type="checkbox"/> 第五場 <input type="checkbox"/> 第三場 <input type="checkbox"/> 第四場 <input type="checkbox"/> 第二場 <input checked="" type="checkbox"/> 男 <input type="checkbox"/> 竹井 <input type="checkbox"/> 朝陽 <input type="checkbox"/> (空白セル)	1166 男 電話いいんですか切っちゃって。 1168 男 輝さん、まだ話している最中でし 1170 男 はい？ 1172 男 名前って？ 1174 男 ああ、輝さん。 1176 男 今、あなたが電話でそう呼んだか 1178 男 えー呼びましたよ。おかしいな。 1182 男 あれ、いいんですか切っちゃって 1184 男 輝さんじゃないんですか電話？ 1186 男 わ、怖いですね。いいんですか、 1188 男 それって絶対に出さないといけな 1190 男 いけませんでしたか？ どなたか
<input type="button" value="OK"/> <input type="button" value="キャンセル"/>	<input type="button" value="元データ"/> <input type="button" value="フィルタ"/> <input type="button" value="+"/>

絞り込んだ後、B列のセリフをコピーして、テキストファイルに貼り付けて読み込ませれば、男のセリフの分析ができる。



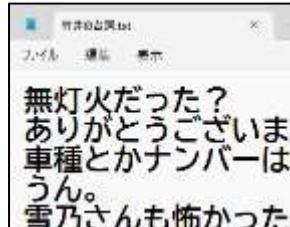
同様に、竹井のセリフだけのテキストファイルを作る。

竹井のセリフに絞り込む



A		ト書き・台詞
1	場・登場人物	ト書き・台詞
12	竹井	無灯火だった?
18	竹井	ありがとうございます。朝陽さんちよ
20	竹井	車種とかナンバーはわからなかつたよ
25	竹井	うん。
27	竹井	雪乃さんも怖かったですね。
33	竹井	あれでも雪乃さんはそこにいらっしゃ
35	竹井	車種もわかんないよね?
38	竹井	見てないんですよね雪乃さんは?
	竹井	あのあたり農道で防犯カメラないんで
55	竹井	んに聞きに行ってもらつてますから。
57	竹井	じゃ私もいいですか。

このB列のセリフをコピーして、新しいテキストファイルに読み込ませて、竹井のセリフだけのテキストファイルを作る。



セリフだけのテキストファイルを User Local に読み込ませる。

User Local のサイト <https://textmining.userlocal.jp/>

にアクセスし、ログインする。

ファイルをアップロードをクリックする



The screenshot shows the User Local AI Text Mining interface. On the left, there is a sidebar with options: '新規作成' (New), '翻訳一覧' (Translation List), 'カスタマイズ' (Customize), '前処理ツール' (Pre-processing Tools), and 'ヘルプ' (Help). The main area has a title '解析したいテキストを入力する' (Input the text to be analyzed). Below it are three tabs: '1つの文書を解析' (Parse one document), '2つの文書を比較' (Compare two documents), and 'Excelクロス集計' (Cross-tabulation in Excel). A red box highlights the 'ファイルをアップロード' (Upload file) button, which is located in the 'フォーム入力' (Form input) section. To the right, there is a note about supported file formats: '以下の形式がアップロードできます。最大200,000文字、10MBまで。' (The following formats can be uploaded. Maximum 200,000 characters, 10MB). The supported formats listed are: テキスト形式(.txt), CSV形式(.csv), Word形式(.docx), Excel形式(.xlsx, .xlsm), PowerPoint Text形式(.pptx), PDF形式(.pdf).

解析するファイルにファイルをドラッグする



The screenshot shows the User Local AI Text Mining interface. At the top, there are three tabs: 'フォーム入力' (Form input), 'ファイルをアップロード' (Upload file), and '履歴から選択' (Select from history). Below the tabs, a note says: '以下の形式がアップロードできます。最大200,000文字、10MBまで。' (The following formats can be uploaded. Maximum 200,000 characters, 10MB). The supported formats are listed as: テキスト形式(.txt), CSV形式(.csv), Word形式(.docx), Excel形式(.xlsx, .xlsm), PowerPoint Text形式(.pptx), PDF形式(.pdf). A text input field '解析対象に名前を設定 (任意)' (Set name for analysis target (optional)) contains '竹井の台詞.txt'. Below it is a '解析するファイル' (File to be analyzed) section, which has a large red box around the 'Drag and Drop File' area. Inside this area, there is a file icon labeled 'TXT'. At the bottom, there are '辞書設定' (Dictionary settings) with a dropdown menu showing 'さらば黄昏', '公開設定' (Publication settings) with '外部に公開' (External publication) selected, and a 'テキストマイニングする' (Text mining) button.

竹井と男の1つの文書を解析の結果。比較をしたいなら、「2つの文書を比較」を使う。

ワードクラウド

竹井

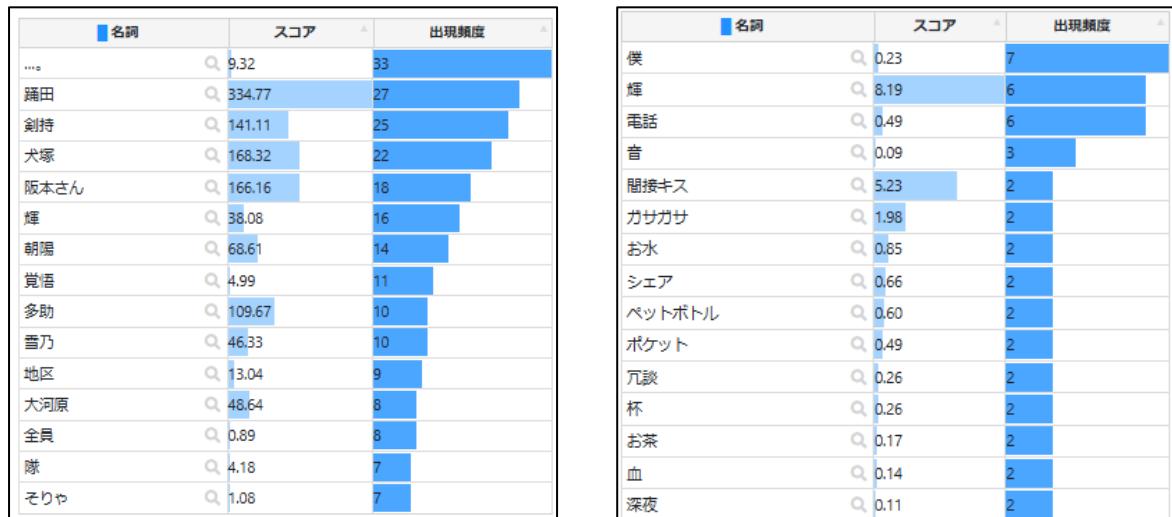
男



单語出現頻度

竹井

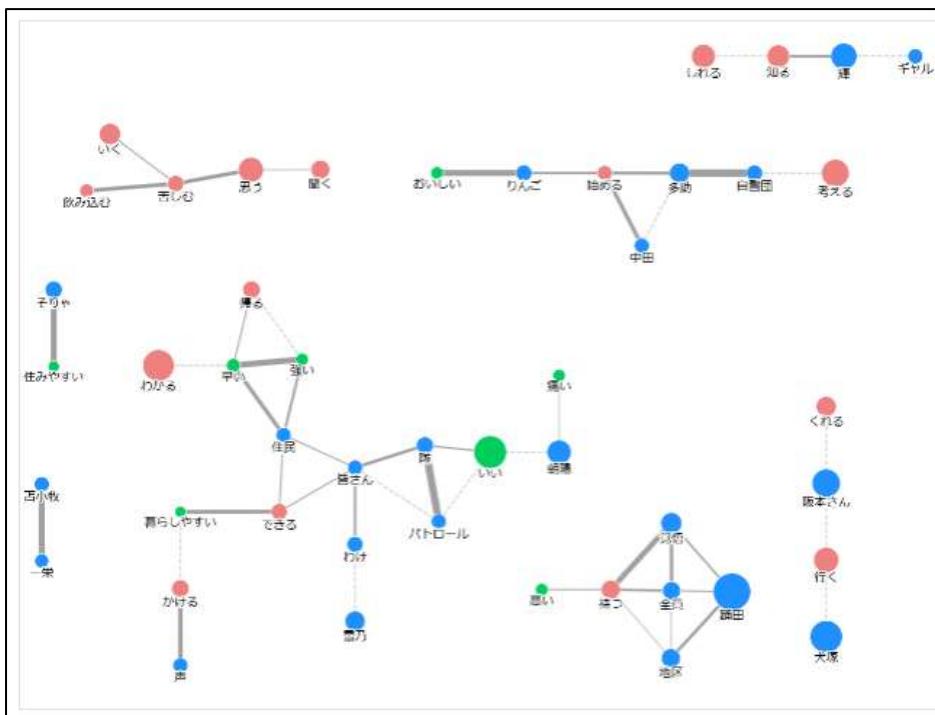
男



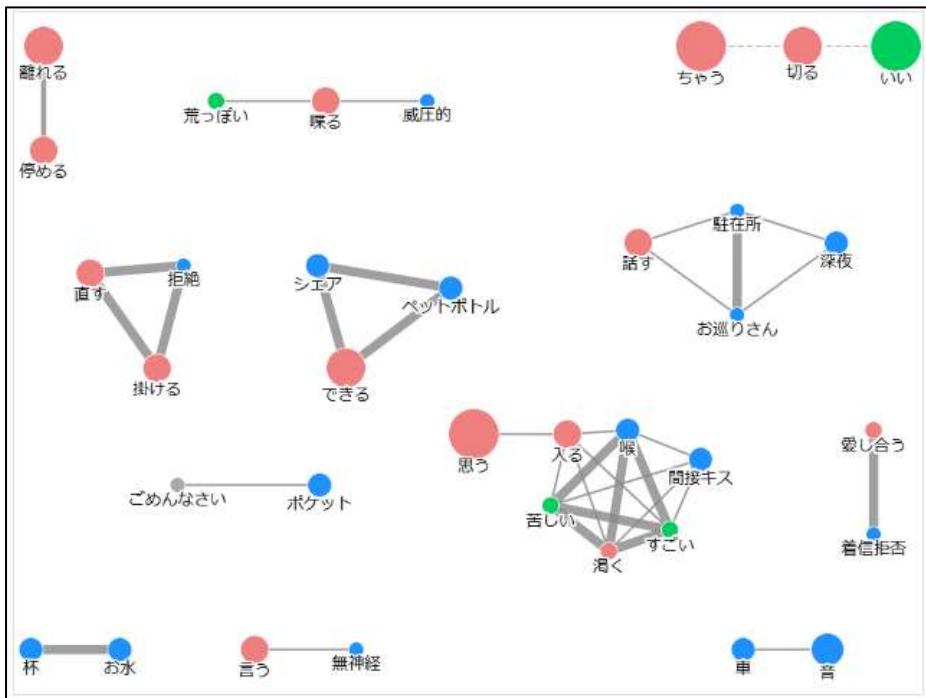
頻出回数が同じ単語でも、スコアが高いものは、その言葉は重要であるとAIが判断しているということ。

共起キーワード

竹井



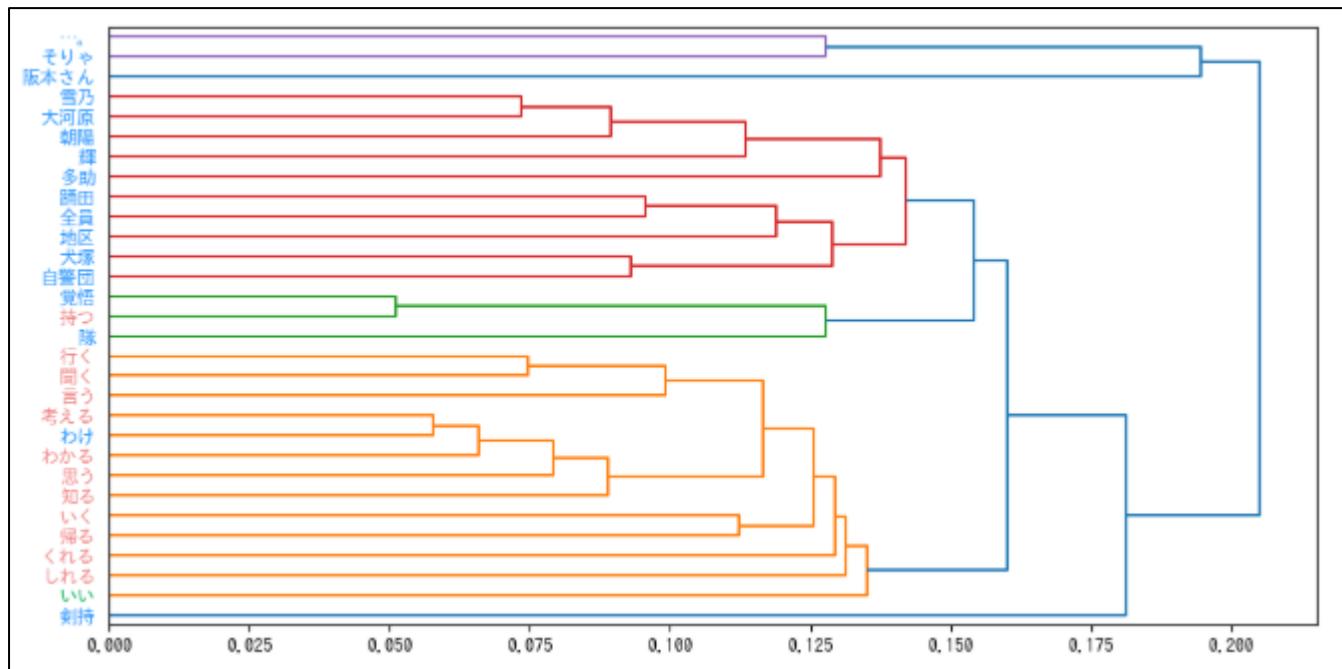
男



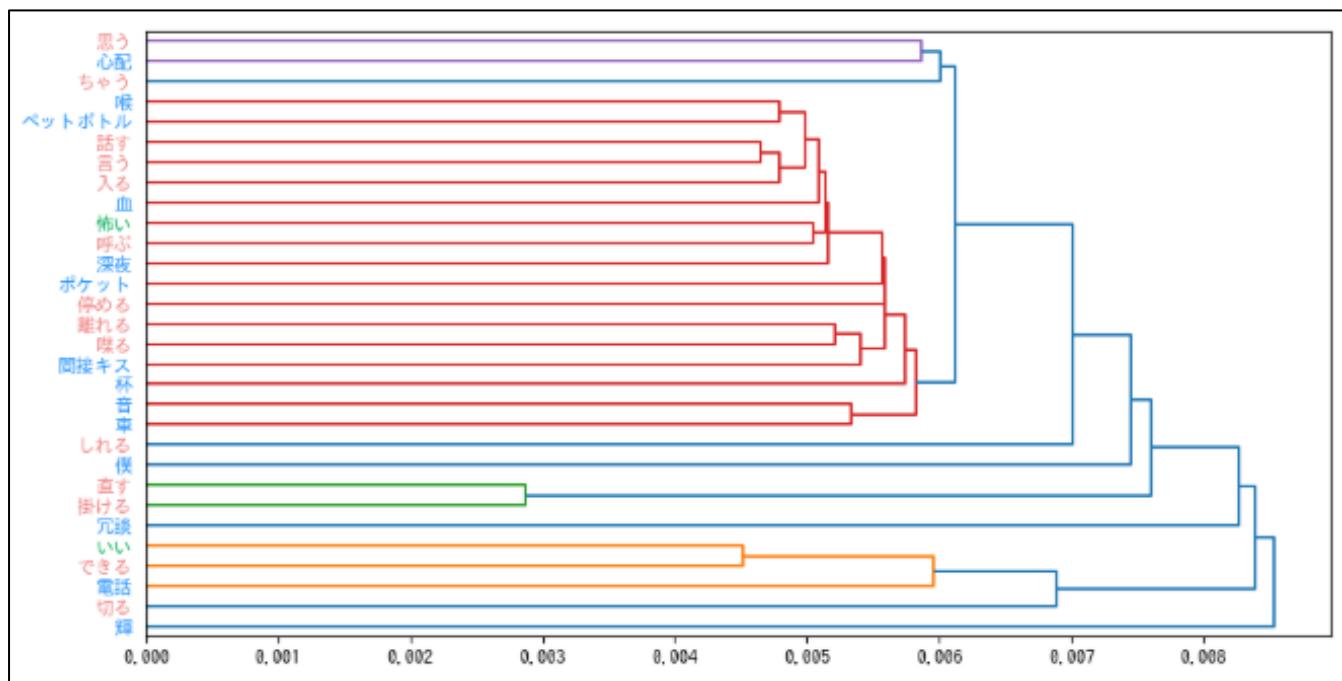
頻出している単語であっても、「繋がり(共起)のパターンがバラバラである」場合は共起ネットワークには出てこない。「輝」は頻出するがと同じ言葉を、文の中で言っていたりしない。あとは、輝の名前だけで繰り返しているなどの原因も考えられる。

階層的クラスタリング

竹井



男



竹井のグループは、公的な義務という軸から来る「固さ」が感じられるが、男のグループの方は「個人的な嫌がらせ」から来るものだろうか、論理的ではない一見関係のなさそうなバラバラな言葉が隣り合っている。

次にやってみることとしては、

二人のデータを合体させて一つのクラスタリングにかけてみたりすると、思わぬ言葉が結びついたりして、戯曲のテーマが浮かび上がったり、二人が危険な領域を共有していることが分かったりする。

「2つの文書を比較」による分析

□ 1つの文書を検討 □ 2つの文書を比較 □ Excel(クロス集計) □ 手動入力

サンプル：複数箇所の高評価・低評価レビューを比較解析

フォーム入力 ファイルをアップロード 既存から選択

以下の形式がアップロードできます。最大200,000文字、10MBまで。
テキスト形式(.txt)、CSV形式(.csv)、Word形式(.docx)、Excel形式(.xlsx)、PowerPoint形式(.pptx)、PDF形式(.pdf)

解析対象に名前を指定（任意）

竹井の山詞(1)

解析するファイル

 TXT

詳細設定を聞く

比較対象に名前を指定（任意）

別の名詞.txt

比較するファイル

 TXT

詳細設定を聞く

文書比較完了手順

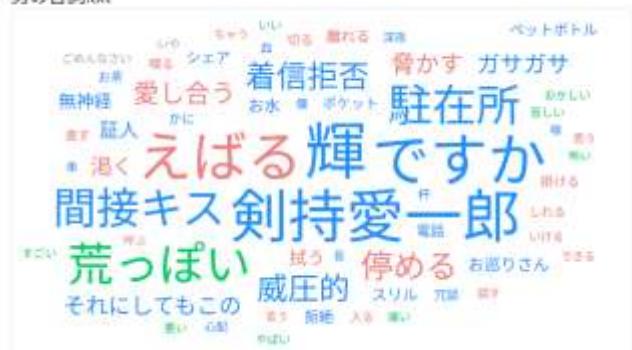
ワードクラウドで比較

スコアが高い単語を複数選び出し、その値に応じた大きさで図示しています。単語の色は品詞の種類で異なっており、青色が名詞、赤色が動詞、緑色が形容詞、灰色が感動詞を表しています。

竹井の台詞



用の台詞 tvt



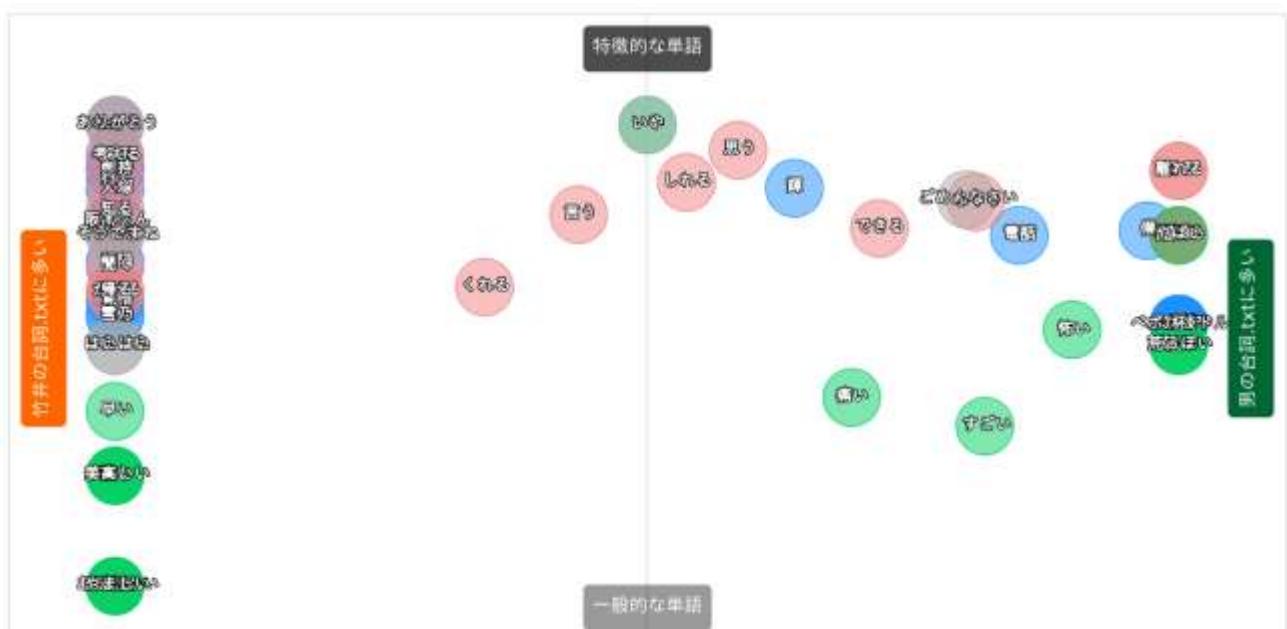
单語分類

2つの文書に出現する単語を、それぞれどちらの文書に偏って出現しているかでグループ分けし、表にしています。グループ中の単語は出現頻度が多い順に並ぶ傾向があります。

竹井の台詞.txtにだけ出現	竹井の台詞.txtによく出る	両方によく出る	男の台詞.txtによく出る	男の台詞.txtにだけ出現
ありがとう 一歩	くれる	いや いい 思う	ごめんなさい ちやう	切る 韋れる やばい
わかる 考える 踊田		しれる 輝 言う	できる 僕 電話	停める 嘴る 掛ける
剣持 行く 大塚		持つ おく	入る いける 呼ぶ	直す お茶 ガサガサ
知る 藤本さん いく			話す 怖い 音 車	シェア ペットボトル
いいえ そうですね			心配 痛い お水	ポケット 冗談 喉
朝陽 聞く かける			すごい	杯 深夜 血
帰る 覚悟 多助				間接キス おかしい
雪乃 苦しむ 地区				悪い 苦しい
いえいえ はいはい				荒っぽい えばる
全員 大河原 そりゃ				よる 乗り越える
隊 りんご わけ				乾く 出会う 切れる
				愛し合う 感じる

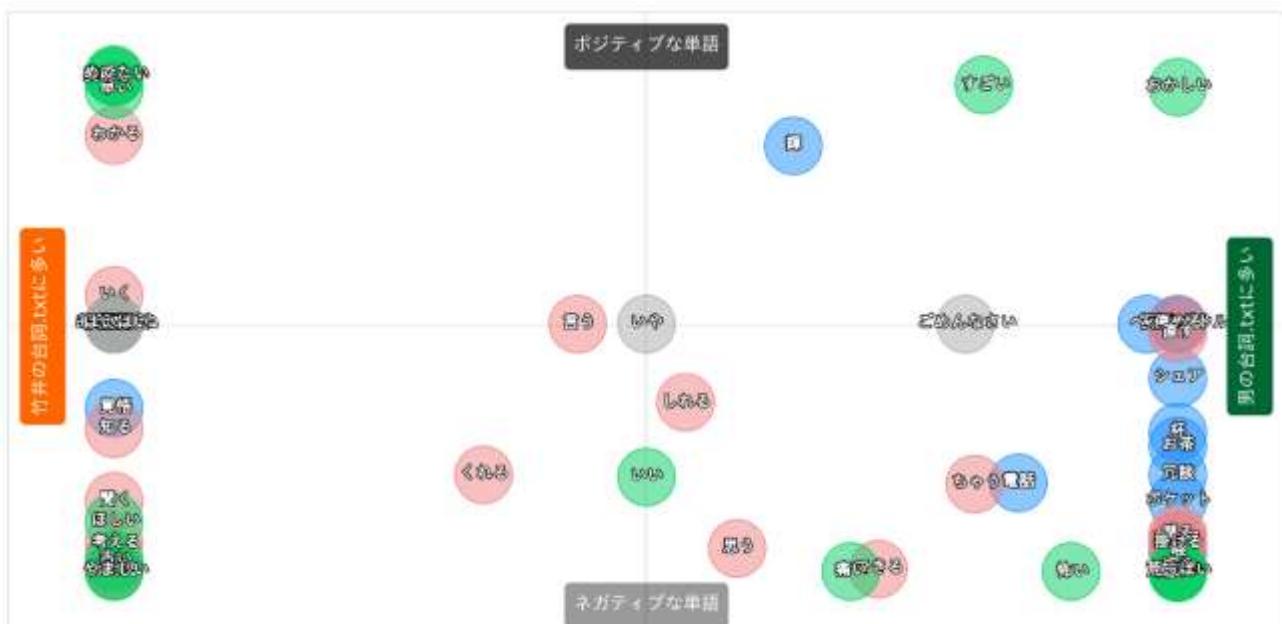
特徴語マップ

それぞれの単語がどちらの文書により多く出現するか、またその単語が文書においてどれだけ特徴的であるかを 2 次元にマッピングしています。上寄りな単語はこの文書の特徴的な単語であることを、下寄りな単語はどんな文書にも出現する一般的な単語であることを示しています。



ネガポジマップ

それぞれの単語がどちらの文書により多く出現するか、またその単語の意味がどれだけポジティブ・ネガティブであるかを 2 次元にマッピングしています。上寄りな単語はその単語がポジティブな意味を持つことを、下寄りな単語はその単語がネガティブな意味を持つことを示しています。



単語の出現比率

2 つの文書から出現回数の多い単語を選び出し、それらが 2 つの文書においてどれぐらいの比率で出現するかをグラフにしています。

