

# 情報検索・探索システムにおけるセレンディピティ Serendipity in Information Retrieval and Search Systems

金津 達也<sup>†</sup>, 吉田香<sup>†</sup>, 古川徹生<sup>†</sup>

Tatsuya Kanatsu, Kaori Yoshida, Tetsuo Furukawa

<sup>†</sup>九州工業大学大学院生命体工学研究科

Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology

kanatsu.tatsuya735@mail.kyutech.jp

## 概要

本研究におけるセレンディピティとは、予期せぬ情報を通じた新たな視点や世界の発見である。本研究では、このセレンディピティを促す情報探索システムを提案する。本システムは、内包的検索と外延的探索の組み合わせとユーザーによる積極的な探索によって、新たな情報に出会う環境を提供する。内包的検索では特定の条件を指定し、外延的探索では直感的に情報を探索する。2つの探索手段の交互利用によって、セレンディピティの促進を目指す。

**キーワード:** セレンディピティ, 拡散好奇心・特殊好奇心, 能動的探索・受動的発見

## 1. はじめに

本研究では、ユーザーが予期せぬ価値ある情報と出会う機会を増やすことを目的とした新しい情報探索システムを提案する。セレンディピティとは、探索中に意図せずして遭遇する有益な情報発見のことを指す。この概念をデジタル情報探索の文脈において再定義し、ユーザーが積極的に探索することで未知の価値ある情報に出会う環境を提供することを目的とする。提案するシステムは、従来の検索エンジンや推薦システムが提供する情報の範囲を超えて、ユーザーに新たな発見や意外な出会いを提供することを可能にする。

従来の情報検索システムや推薦システムでは、ユーザーの過去の行動や明示的なクエリに基づいて情報を提供するため、予期せぬ新しい情報との出会いが限られていた。本研究では、この問題を解決するために、内包的検索と外延的探索を融合させた新しい情報探索手法を導入した。内包的検索では、ユーザーが特定の条件を指定して情報を絞り込む一方で、外延的探索では、ユーザーが直感的に情報を探索し、計画されていない情報発見の可能性を広げることを可能にする。

本研究のキーアイデアは、内包的な情報検索と、外延的な情報探索を組み合わせ、対話的かつ視覚的な情報探索のスキームである。提案システムでは、探索

対象がマッピングされた2つの空間が示される。第一のマップは、キーワードなど言語的に表現可能な空間である。この空間では、ユーザーが探索目的に合った領域を意識的に絞り込むことができる。これを本研究では「内包的検索」と呼ぶ。一方、第二のマップは、第一のマップに対して統計的に独立な、直交した空間である。ここでは、言語化した空間と独立な方向へ、ユーザーは新たなアイテムを求めて探し求めることができる。これを本研究では「外延的探索」と呼ぶ。

この2つの空間を交互に訪れることで、すなわち、言語化できる関連性と言語化できない関連性による手がかりを交互に用いることで、セレンディピティに出会う機会を促進する情報探索システムを実現することが本研究の狙いである。

## 2. セレンディピティ

### 2.1 Explicit style map と Implicit aesthetic map

「セレンディピティ」という言葉は通常、「偶然に起こった幸運な発見や出来事」と定義される。科学研究の文脈では、思いがけない現象から画期的な発見をした際に「セレンディピティ」という語が使われており、ペニシリンの発見が代表的事例として挙げられる。近年では自然科学の領域にとどまらず、ビジネスなどのマネジメント領域や、推薦システムなどのデータサイエンス領域でもセレンディピティの概念が取り入れられている。これらの領域では、組織マネジメントにおける新たな価値創造を支援することや [1], 推薦システムのユーザー満足度を向上することを目指している [3]。このように、セレンディピティに出会う機会を増やすことは、さまざまな分野で重要な関心事となっている。本論文では、セレンディピタスな情報に出会う機会を促進する情報検索・情報探索システムについて焦点を当てる。すなわちユーザーが目的とする情報の検索・探索をする行動を通して、ユーザーにとって本来目的外の有意義な情報に遭遇し、そこを起点に新

たな分野や領域を知る機会を促進する情報システムのフレームワーク提案を目的とする。

従来、「セレンディピティ」はさまざまな分野でそれぞれに異なる定義されてきた。推薦システムの分野では、セレンディピティとは「関連性」「新奇性」「意外性」の三要素を持つものであるとししばしば定義される [4]。すなわち、ユーザーの関心事に関連がある新奇なもので、かつそれがユーザーにとって意外なものがセレンディピティであり、このようなアイテムをいかにして見つけてユーザーに推薦するかが重要とされている。一方、ビジネスの分野ではセレンディピティを「行動主体」「驚き」「価値」の三要素から成ると定義される [2]。ここで「行動主体」とは、セレンディピティなイノベーションやマーケットを見つけようと能動的に行動する個人や組織であり、「驚き」とは発見されたものへの新奇性や意外性、「価値」は発見されたイノベーションやマーケットが生み出す価値である。すなわち行動主体と発見対象との関係性の枠組みで理解されている。また自然科学の分野では、ペニシリンの発見のように、意図しない出来事に対する気付きから大きな発明や発見につながる現象として理解されている。このようにセレンディピティの定義は分野ごとに異なる。さらに同じ分野であっても研究者ごとにセレンディピティの定義が異なっており、統一された定義が存在しないのが実情である。そこで本論文では、セレンディピティの原義に立ち戻り、その意味を再定義することから試みる。

「セレンディピティ」という言葉は、ペルシャの民話「セレンディップの三人の王子」という物語に由来する。この民話はセレンディップ国の三人の王子が世界を旅する物語であり、王子たちはある目的を持って旅をするが、その途中で出会うさまざまな目的外の出来事に対しても積極的に関心を持って観察し、そこから得られた発見によって思いがけない幸運を手にする。ホレス・ウォルポールはこの物語を書簡の中で紹介し、「王子たちが偶然と賢さによって自分たちが求めていなかったものを常に発見している」と述べ、このような発見を「セレンディピティ」と名付けた [5]。

このホレス・ウォルポールが名付けたセレンディピティの原義は、他分野でのセレンディピティの定義と共通点が多いものの、若干の相違もある。そこで本論文では、ホレス・ウォルポールによる原義をふまえて、「能動的な探索行動によって、新奇で価値のある情報を発見すること」と再定義する。この再定義を要素に分解すれば、「能動的行動」「新奇性」「価値」となる。これはビジネス分野における三要素とほぼ重なる。す

なわち、本論文におけるユーザーは「能動的な情報探索行動」を行うことで「新奇性」のあるものを発見し、そこから新たな視点や新たな世界という「価値」を見出す存在であるとみなす。またここで言う「価値」とは、その情報をもたらす新たな視点や新たな世界である。すなわち本研究では、セレンディピティの価値を、発見したアイテム自体ではなく、その背後にある新たな世界に対する価値として定義する。たとえるなら、「アリスの白うさぎ」の価値は白うさぎ自身ではなく、「白うさぎが連れていく新しい世界」の価値として評価される。

上記の定義に基づくと、セレンディピティ遭遇までのプロセスは以下のようにモデル化できる。

1. ユーザーが当初意図していた情報（一次標的）の探索フェーズ。このフェーズは一次標的に特化した好奇心（特殊好奇心）により駆動される。このフェーズでは、多くの情報から目的の情報を絞り込むフィルタリングが行われるため、ユーザーは「情報選択的不注意」の状態に置かれている。
2. 一次標的の探索を通してさまざまな情報に触れることで、一次標的以外の情報への好奇心（発散好奇心）が活性化するフェーズ。これにより、「情報選択的不注意」の状態からユーザーを解放する。
3. 発散好奇心が活性化することにより、一次標的だけでなく、そこから他の関連情報にも関心が生じる、派生的探索フェーズ。このフェーズは発散好奇心により駆動される。
4. 派生的探索により、新奇な情報に出会い、ユーザーの関心が新奇情報にフォーカスされるフェーズ。
5. 新奇な情報から提示される新たな視点や新たな世界に気づき、ユーザーがセレンディピティを「発見した」と感じるフェーズ。

このようなユーザーの能動的探索行動を支援することでセレンディピティを促進する情報探索システムを実現することが本研究の目的である。より正確に言えば、そのような情報探索システムを構築することを通して、セレンディピタスな情報システムにはどのような要素が必要かを明らかにし、セレンディピタス情報探索システムが持つべきフレームワークを明らかにすることが本研究の目的である。

### 3. 提案手法

第2節では、本論文におけるセレンディピティを「能動的な探索行動によって、新奇で価値のある情報（アイテム）に出会うこと」と定義した。すなわち「能

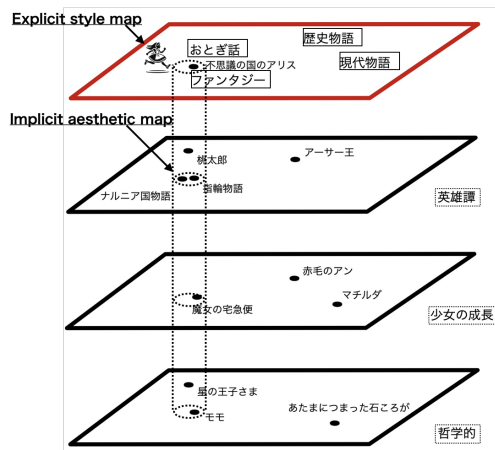


図 1: アリスのウサギ穴の図：探索対象である書籍全体が 3 次元空間にマッピングされている。3 次元空間における水平面は、例えばおとぎ話やファンタジーといったジャンルやスタイルのような言語で表現可能な空間軸である。他方の、3 次元空間における垂直軸は、例えばナンセンスや少女の成長といった雰囲気や印象などの言語で表現困難な空間軸である。図中では例となるキーワードを双方の軸に表示しているが、本来は垂直軸にはキーワードは存在しない。水平面と垂直軸を交互に探索していくことで、セレンディピティに出会う機会を促進する狙いがある。

動的行動「新奇性」「価値」である。またここでいう「アイテムの価値」とは、「アリスの白うさぎ」、すなわち新たな視点や新たな世界に気づききっかけを与えることであると定義した。この定義に基づく、セレンディピティ探索システムでは、以下の要素が必要になる。

**能動的行動** ユーザーの能動的探索行動をアフォードする探索インターフェースを持つ。身体感覚的に新たなアイテムを探索しやすく、能動的な探索を可能とする。

**新奇性** 探索によって、探索目的以外の新奇なアイテムに出会えるようにする。しかも探索目的となんらかの関連性を持ちつつも、意外性のある新奇アイテムが探索で見つけられることを支援する。

**価値** 新奇なアイテムに出会った時、その背後にある新たな視点や新たな世界を提示でき、ユーザーがその価値を発見することを支援する。

これらを実現するため、本研究では視覚的探索システムを提案する。提案システムでは、探索対象のアイテムは 4 次元空間に埋め込まれている。そしてこの空間は直交する 2 つの空間に分解表現される。すなわち、2 つの空間の点の組み合わせで、4 次元空間の 1

点を指定することができる。

2 つの直交空間のうち、一方は言語化可能なジャンルやスタイルによりアイテムをマッピングした空間である。われわれはこの空間を 'Explicit style space' と呼び、この空間にアイテムをマッピングして地図として可視化したものを 'Explicit style map' と呼ぶ。ユーザーに対しては、ジャンルやスタイルなど、言語化可能な分類に基づく領域が表示されたアイテムのマップとして提示される。'Explicit style map' では、ユーザーは自分の探索標的を意識的に探索することができる。

もう一方の空間は、その情報やアイテムが持つ高次の意味的類似性によりアイテムをマッピングした空間である。われわれはこの空間を 'Implicit aesthetic space' と呼び、その空間にアイテムをマッピングして可視化したものを 'Implicit aesthetic map' と呼ぶ。この空間では、作品の雰囲気など言語化できない類似性に基づきアイテムが配置されている。このマップ内でユーザーは言語化された意識的探索ができないため、確率的探索な探索を試行したり、無意識的な好みに駆動された探索をしたりすることになる。すなわち好奇心駆動型の探索行動することが求められる。

図 1 は、このフレームワークでユーザーがどのようにアイテムを探索し、セレンディピティに遭遇するかを示したものである。この図では、ユーザー「アリス」が新たな文学の世界に出会うプロセスを描いている。彼女は「不思議の国のアリス」を読み、ファンタジーの世界に興味を持つ。そこで、最初は Explicit style map を探索して、ファンタジー小説の領域にたどりつく。Explicit style map ではジャンルなどのキーワードが示されているため、アリスは目的の領域に容易にたどり着ける。なおこの段階で、アリスはファンタジー小説以外には関心を持っていない。ファンタジー小説の領域を見つけたアリスは、Explicit style map の該当領域の一点を、Target of Interest (ToI) として指定する。Explicit style map 上で ToI は 1 点だが、あたかもうさぎの穴のような奥行きを持っている。この奥行きが、Explicit style space と直交する Implicit aesthetic space である (Implicit aesthetic space も 2 次元であるが、図では 1 次元として描いている)。Implicit aesthetic map には、Explicit style map の ToI (ファンタジー小説) に属するアイテム集合、すなわちさまざまなファンタジー小説のさまざまな本がマップされている。Implicit aesthetic map 上では、似た雰囲気を持つアイテムが近くに配置されており、さまざまなファンタジー小説を求めて探索する

ことができる。たとえば「指輪物語」の近くには「ナルニア国ものがたり」があるだろうし、「トムは真夜中の庭で」の近くには「思い出のマーニー」があるだろう。ここでアリスが「魔女の宅急便」が気に入り、それを Implicit aesthetic map の ToI として指定したとする。すると今度は、その本と似た雰囲気を持つさまざまなジャンルの本が、Explicit style map の上に表示される。そこには、「魔女の宅急便」と似た雰囲気を持つ他のジャンルの本が可視化されている。たとえば同じく少女の成長譚である「赤毛のアン」をアリスは見出すかもしれない。同様に、哲学的な雰囲気のある「モモ」を見つけたアリスは、同じく哲学的な雰囲気を持つ「星の王子さま」を見出すかもしれない。この見出された「赤毛のアン」や「星の王子さま」はファンタジーの枠から外れており、まさに新たな世界を提示する「アリスの白ウサギ」となっている。

Explicit style map では、マップの意味を言語的に捉えることができるため、ユーザーは関心の領域を絞り込みやすい。われわれはこれを「内包的探索」と呼ぶ。一方、Implicit aesthetic map では、マップの意味を言語的に捉え難いため、ユーザーは関心の領域を求めて好奇心駆動的な探索をする行う必要がある。われわれはこれを「外延的探索」と呼ぶ。Explicit style map における内包的探索による意識的絞り込みと、Implicit aesthetic map における外延的探索による好奇心的発見行動の交互的探索が本研究のキーアイデアである。ここで Explicit style map と Implicit aesthetic map を交互に利用する探索は、2 節で述べたセレンディピティに至るプロセスと一致することに注意してほしい。またこの探索方法は、セレンディピティの三要素「能動的行動=主体的な探索行動」「新奇性=一次標的以外との遭遇」「価値=新たな世界の提示」を満たしている。

### 3.1 アイテム埋め込みとモデリング

本研究では、探索の対象であるアイテムを、Explicit style space  $\mathcal{U}$  と Implicit aesthetic space  $\mathcal{V}$  の積空間  $\mathcal{U} \times \mathcal{V}$  に埋め込む方法を提案する。アイテム  $i$  の座標を Explicit style space で  $u_i$ 、Implicit aesthetic space で  $v_i$  とし、これを  $(u_i, v_i) \in \mathcal{U} \times \mathcal{V}$  にマッピングすることで、アイテムの特性をモデリングする。

アイテムの高次元特徴ベクトル  $\mathbf{x}_i$  は、写像  $f: \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{F} \subseteq \mathcal{X}$  を通じて生成されたと考え、この写像  $f$  により、 $\mathcal{F}$  が  $\mathcal{U}$  を底空間とするファイバー束になるとする。さらに、アイテムのジャンルやタグなどの補助

情報  $\mathbf{y}_i$  を利用して、Explicit style space への埋め込み  $u_i$  を推定し、主情報  $\mathbf{x}_i$  を用いて Implicit aesthetic space への埋め込み  $v_i$  を推定する。

情報検索・探索の対象となるアイテムの数を  $I$  個とし、それぞれのアイテムが持つ主情報  $\mathbf{x}_i$  と補助情報  $\mathbf{y}_i$  を  $\mathcal{X} = \mathbb{R}^{d_x}$  および  $\mathcal{Y} = \mathbb{R}^{d_y}$  のベクトルとして表現する。主情報はアイテムの美的特性を示し、補助情報は明確な分類情報を提供する。

これらのデータから、アイテムのファイバー束への埋め込み  $(u_i, v_i)$  と滑らかな埋め込み写像  $f: \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{X}$ ,  $g: \mathcal{U} \rightarrow \mathcal{Y}$  を推定することが目的となる。推定方法としては、アイテムを補助情報  $\mathbf{y}_i$  を用いて explicit style space に、主情報  $\mathbf{x}_i$  を用いて implicit aesthetic space に埋め込む二段階アプローチを採用する。推定では、以下の目的関数を最小化する。ここで  $R(u)$  は  $u$  の事前分布に相当する正則項である。

$$E_1[U] = \frac{1}{2} \sum_i \|\mathbf{y}_i - g(u | U)\|^2 + \sum_i R(u_i)$$

$$E_2[V] = \frac{1}{2} \sum_i \|\mathbf{x}_i - f(u_i, v_i | U, V)\|^2 + \sum_i R(v_i)$$

このアプローチにより、高次元データの効率的な低次元表現と直感的な可視化が可能となる。

## 4. おわりに

本論文では、セレンディピティの定義から再考し、再定義したセレンディピティの発生を支援するシステムの機能：内包的検索・外延的探索に基づいて、モデルを提案・実装した。今後の研究では、一般の被験者を対象に実証実験を行うことを考えている。

## 参考文献

- [1] Marco Balzano, (2022) “Serendipity in management studies: a literature review and future research directions”, *Management Decision*, Vol. 60, Issue. 13, pp. 130-152.
- [2] Christian Busch, (2022) “Towards a Theory of Serendipity: A Systematic Review and Conceptualization”, *Journal of Management Studies*.
- [3] Li Chen and Yonghua Yang and Ningxia Wang and Keping Yang and Quan Yuan, (2019) “How serendipity improves user satisfaction with recommendations? A large-scale user evaluation”, *The world wide web conference*, pp. 240-250.
- [4] Denis Kotkov and Shuaiqiang Wang and Jari Veijalainen, (2016) “A Survey of Serendipity in Recommender Systems”, *Knowledge-Based Systems*, Vol. 111.
- [5] Horace Walpole, (1960) *The Yale Editions of Horace Walpole's Correspondence*, Vol. 20, pp. 407-411.