

ウェブを介した情報発信における 提供者と提供物の特性を考慮したページ・リンク生成過程の記述

Description of page and link generation processes
considering the characteristics of information provider and contents in WWW

熊澤 健人[†]
Kento Kumazawa

中平 勝子[†]
Katsuko T. Nakahira

北島 宗雄[†]
Muneo Kitajima

1 はじめに

近年の急激な ICT の発達によって、ウェブコンテンツの閲覧・利用は一般に広く浸透した。加えて、コンテンツの作成やウェブリンクなどによるネットワーク形成に関してもユーザ個人の技術力に大きく依存していた頃と比べ、ユーザの技術力をサポートするソフトウェアなどの普及によって、ユーザ個人の技術力への依存度が低下し、個人による情報発信が盛んになってきている。

ユーザがウェブを介した情報発信を行う場合、表現するコンテンツを作成し、ウェブに公開する。公開されたコンテンツは、多くのユーザの手によって他のコンテンツとリンクが形成される。このようにして公開されたコンテンツはウェブに拡散していくと考えられる。上記のようにウェブにおける情報発信・拡散は、人 - コンテンツ間の相互作用を軸になされる。

そこで本稿では、ウェブにおける情報発信活動を理解するために、個人によるウェブコンテンツ・ウェブリンク生成過程モデルの構築を目指す。

この過程はウェブコンテンツ・ウェブリンク生成に携わるユーザ個人の持つ ICT リテラシーや情報発信意欲などの内的要因や、ICT リテラシー教育や情報発信活動に割り当てることができる時間の制約などの外的要因による個人の特性に影響されると考えた。こういった特性を有したウェブを作る人々を取り巻く環境の変化などがウェブリンク構造に与える影響について考察する。

以下の章では、コンテンツ生成・ウェブ空間成長プロセスモデルの記述を行った後、ウェブにおけるネットワーク構造の特徴について述べる。最後に生成者・生成物特性を考慮したシミュレーションを行いその結果から、特定の特性がウェブの成長やネットワーク構造に与える影響について考察を行う。

2 ウェブ空間成長過程の記述

ウェブ空間成長過程とは、『生成者』『生成物』『環境』が複雑に絡み合うことによって、ウェブ空間のコンテンツ数の増加やコンテンツ間のネットワークが形成される過程である [1]。図 1 はウェブ空間成長過程を模式的に示したモデルである。

本モデルでは、生成者と生成物に加えて、生成者を取り巻く環境、生成物の集合からなるウェブ、生成者の活動、の 5 層に分けて記述した。生成者・生成物層では、それぞれ考えられる固有の特性を示し、環境層には、生成者特性や生成者の活動に影響を与えると考えられる特性を記述した。生成者活動層には、生成者の情報獲得プロセスと、情報発信・拡散プロセスを

記述した。

生成者がウェブを介して情報収集を行う場合、はじめにウェブコンテンツを閲覧する。閲覧によって知覚した情報を生成者の既有知識や言語能力を用いて理解をし、情報獲得を行う。この際、生成者の情報獲得量と情報獲得に要する時間は、個人の理解力や情報探索能力などに依存する。生成者の情報発信・拡散活動では、既有知識を用いてコンテンツの、生成・削除・更新・公開やリンクの生成を行う。既有知識でこれらが行えない場合は、知識獲得のために作業を中断し、あらゆる手段で問題解決のための知識獲得を行う。

生成者は、個々で特有の特性を有している。この特性は、個人特性の多くは周囲の環境によって変化する可能性がある。例えば、住んでいる国によって使用可能な言語が異なる、生成者の職業によって資本が異なる、他者からの教育によって ICT リテラシーが異なるといった具合である。さらに生成物の持つ特性には、生成者や環境の特性に依存するものもある。

本稿では、ウェブ空間成長の根幹となっている生成者固有の特性やそれに影響を与える環境特性によってウェブ空間の成長は変容すると考えている。

3 コンテンツ生成過程の記述

コンテンツ生成過程とは、ウェブ空間成長過程の根幹を成す過程であるウェブを介した情報発信を行う際のコンテンツ生成の過程を指す。このモデルはその過程を『意欲』『制作』『発信』の三階層に分けて表現したものである。

このモデルでは、人の情報発信を行いたいという意欲がコンテンツ生成の起点となる。コンテンツ生成活動が始まると、生成者はコンテンツの設計・制作・編集を行う。ウェブコンテンツ制作において、作業効率には生成者のリテラシーに依存する。そこで、ウェブコンテンツ制作活動の状態を“Working”と“Idling”の二種類に分けて考える。Idling 状態では、コンテンツ制作に必要な知識の獲得や、技術の習得を行う。Working 状態では、新たなコンテンツの制作や、既存コンテンツの編集などを行う。コンテンツ制作では、Working 状態と Idling 状態が繰り返し行われる。Idling 状態では、制作活動が中断され、設計を実現するための知識や技術の獲得を行う。そのため、Idling 状態が長い生成者はコンテンツ制作効率低下し、Idling 状態の短い生成者は、コンテンツ制作効率向上する。

ICT リテラシーの高いユーザは、概念設計を実現するための知識を多く有しているため、知識・情報獲得のための Idling 状態が短くなる。よって、ICT リテラシーの高い生成者は、低いユーザに比べ、同程度のコンテンツを制作する際に要する時間が短時間で済み、コンテンツ生成サイクルの周期が短くなるこ

[†] 長岡技術科学大学

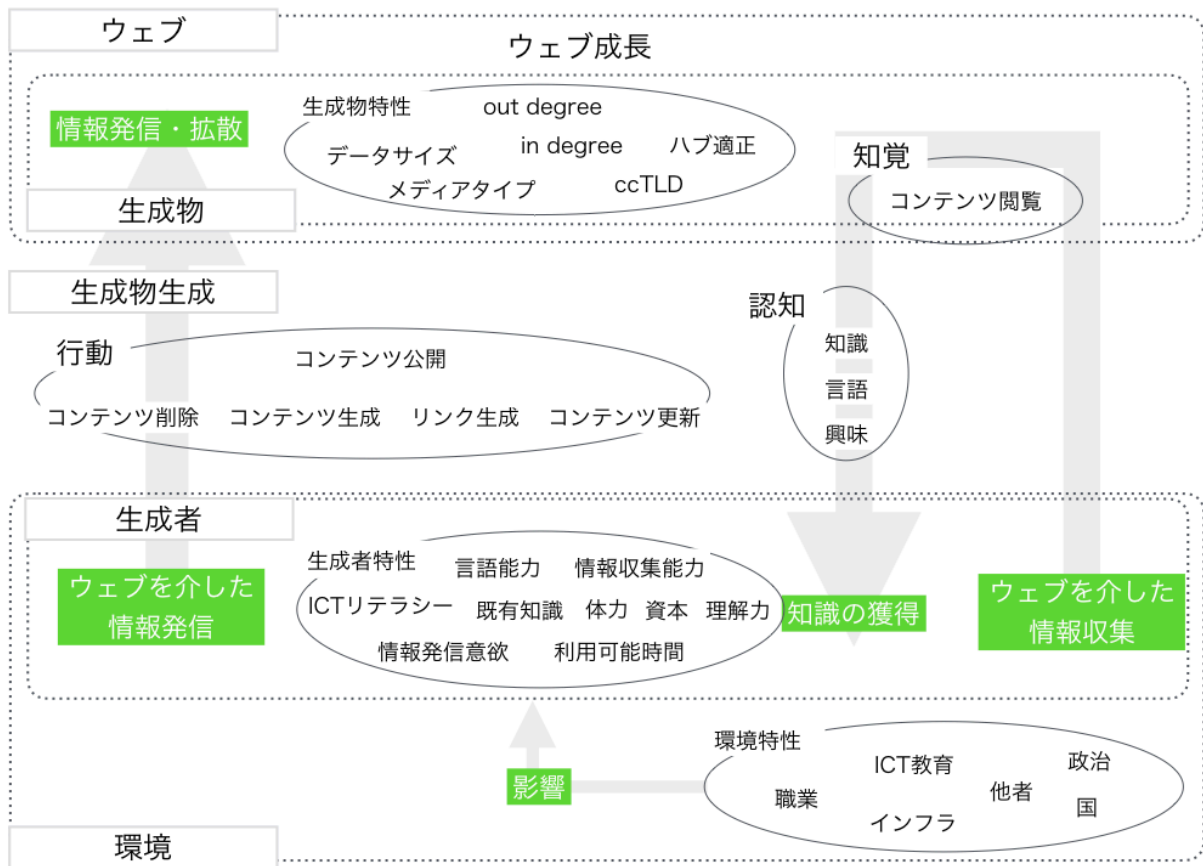


図1 ウェブ空間成長プロセス

とで高頻度で活発な情報発信活動を行うと考えられる。本章では、ウェブ空間成長プロセスモデルに基づいて生成者特性・環境特性と、生成物特性について本稿のシミュレーション実験で用いる特性を中心に記述する。

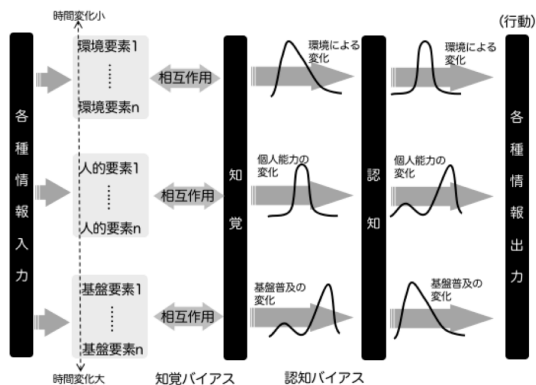


図2 e ネットワークを元にした、人の認知特性にもとづいた情報獲得 (情報入力) から情報発信 (情報出力) の流れ

図2に人の認知特性に基づいた情報獲得から情報発信の一連の流れを示す。人は外界から各種情報をオブジェクトとして獲得する(情報入力)。これは、知覚に相当する行為である。しかし人はオブジェクトを知覚する際、自らを取り巻く様々な要素

のバイアスがかかった形で情報を知覚する。バイアスがかかった知覚情報は、人の脳内で様々な形に変換され認知される。認知されたオブジェクトは、これら過程において、個人が置かれた環境や情報基盤、個人が持つ素質としての人的要素によって、さらに認知バイアスがかかる形で情報として脳内に蓄積される。その情報が、人にとって「発信したい」と思える情報であった場合、情報発信(情報出力)という行動を取る。その際にも個人が置かれた環境や情報基盤、素質といった要素によるバイアスがかかる形で出力される。図2において、様々な箇所に環境要素・基盤要素・人的要素に由来するバイアスがかかるため、外界に表出する情報は、元々人が享受した情報(オブジェクト)とは異なる形で表現されることもしばしばある。以上のプロセスをシミュレーションとして実装する際、各種バイアスとして置いた要素と人の関係は、「相互作用」として表現される。また、これは単体の人についての図式であるが、実際には人が孤立してウェブ空間で情報獲得・情報発信を行うことはなく、多くの人が発信する情報に影響される形で情報発信を行う。そのため、情報を享受する人が属するコミュニティやどのような基盤・人的要素を持つのか。コミュニティ内に同様の環境・基盤・人的要素を持つものがどの程度存在するか。これらによって人の情報発信行動、すなわちウェブ空間成長に大きな影響を与える。

生成者特性

生成者特性には、内的要因によるものと外的要因によるものが存在している。内的要因とは、技術力や言語能力、情報の発信や生成意欲といった生成者固有の能力や心的特徴を指す。外的要因とは、生成者の住む地域や職業など生成者を取り巻く環境特性を指す。内的要因による特性は、しばしば外的要因である環境特性の影響を受けて変化する。

内的要因

【ICT リテラシーレベル】情報リテラシー、コンピュータリテラシー、メディアリテラシーなど様々な情報通信技術を扱う能力を指す。コンテンツ生成過程にて、生成能力や生成効率にもっとも関連すると考えられる特性である。ICT リテラシーの高さによって、コンテンツ生成過程における Idling 状態にある時間が変化する。ICT リテラシーが低い場合は、設計したコンテンツを実現するためにより高頻度の情報収集や学習を行う必要があり Idling 状態になりやすい。ICT リテラシーは、学習や、教育、経験によって向上する。

【情報発信意欲】ICT を用いて情報の発信・拡散を行おうとする意欲を指す。コンテンツ生成過程において、コンテンツ生成の起点となる特性である。情報発信意欲が高い生成者は、より多くのコンテンツ生成・発信を行おうとする。そのため、この特性によって生成者のコンテンツ生成頻度が変化する。この特性は、生成者の持つ環境特性や近傍の生成者など外的要因によって変動しやすいと考える。

【利用可能時間】生成者が情報発信や、ICT に関する学習、生成したいコンテンツについての学習に対して費やすことができる時間量を指す。この特性は、生成者特性である体力や、環境特性である職業などによって決定される。生成者のコンテンツ生成に対する、没頭度合いによっては本来割り当てることが可能な時間を超えて、コンテンツ生成を行う場合などがある。職業が変化した場合などは急激にこの特性の値が変動する場合がある。

【情報収集能力】生成者が Idling 状態にある時など、学習を行っている際に求めている情報を収集する能力である。この能力が高いと Idling 状態で、情報収集にあてる時間が減少するため Idling 状態の総時間が減少する。この能力は、経験や教育などによって上昇すると考えられる。

【理解力】生成者が教育を受けた際や学習を行なった際に、内容理解を行う能力である。理解力が高い場合、生成者が Idling 状態で学習に費やす時間が減少する。この特性は、生成者が潜在的に持っている特性であり他の特性比べて変動しにくい。

外的要因

【国 / 政治】生成者の在住している国を指す。国によって利用言語や、情報技術の発展度、インフラ整備度、ICT 教育のレベルなど生成者特性の変化に関連すると考えられる環境特性が決定する大きな要因の一つである。

【ICT リテラシー教育】ICT リテラシー教育の質の差によって、ユーザの持つ ICT リテラシーの成長度に違いが現れる。ユーザの成長に適した教育が行われている場合、その教育を受けたユーザの ICT リテラシーレベルが上昇する。逆に教育の質が悪いとユーザの ICT リテラシーレベルは上昇しにくい。

生成物特性

生成物特性は、コンテンツの理解に関わる特性が多い。本稿ではコンテンツの特性は一定とするため、情報発信の活発度を確認するための特性のみ説明する。

【outdegree / indegree】ウェブのネットワーク構造を確認するための特性である。コンテンツの他のコンテンツとの接続状況を把握することができる。

4 ウェブにおけるネットワーク構造の性質

これまでのネットワーク解析によって、ウェブページとウェブリンクから形成されるネットワーク構造はスケールフリー性やスモールワールド性といったネットワーク構造の特徴を持つことが指摘されている [2][3]。スケールフリー性とは、一部のノードが大量のリンクを持ちそれに対して、大部分のノードは少数のリンクしか持たないようなネットワーク構造のことである。さらに近年では、上記以外の構造的特徴も確認されている。ウェブの発達によって、ウェブコンテンツの商業利用などが多くなり、機械的に多くのリンクを持つウェブコンテンツが大量に生成されるようになった。これによって、べき的減衰構造から外れた五月雨的構造が現れると考えられている。五月雨的構造の発現は、機械的リンク生成が大きく関すると考えられているため、人の営みにのみ着目した場合は発現しないことが予想される [4][5]。

本研究では、ウェブにおけるネットワーク構造特徴の観測・分析を行うために outdegree 図を用いてネットワーク構造を表現する。outdegree とは、グラフ理論における用語であり、有向グラフにおいて 1 つの頂点から出て行く枝の数を表す。ウェブ空間を巨大な有向グラフと見なせば、これは 1 つのウェブページ上から他のウェブページに対して張られているリンクの総数を表すものである。本研究で用いる outdegree 図とは、横軸にウェブコンテンツが有する outdegree、縦軸にその outdegree を持つウェブコンテンツ数をとる両軸対数グラフのことである。

石原ら [6] によって outdegree 図を国の総ページ数の尺度に沿って整列したとき、outdegree 図は、ウェブページの総数が増加するにつれて、べき乗則に従う分布へと近く傾向が示されている。これは、共時的な分析ではあるが一国について通時的に追跡してその成長を確認するためには長い時間を要してしまう。そのため成長過程を確認するためには、ウェブ空間の成長をモデル化し再現する必要がある。

5 実験結果/考察

本稿では、生成者・生成物によるマルチエージェントシミュレーションを用いてウェブ空間成長の分析を行う。ICT リテラシーの初期値が等しい、30 程度の生成者群が存在するコミュニティを作成する。このコミュニティ内で行われる情報発信活動を再現・観察する。

コミュニティに属する生成者群の ICT リテラシーの違いによって生じる、情報発信の活性度の違いを観察するために、生成者群の ICT リテラシー以外の特性値を一定とする。

ICT リテラシーは、教育・学習・経験などによって変化する特性値である。そこでコミュニティにおける ICT リテラシー教育の質が高いコミュニティ 1 と、ICT リテラシー教育の質が低いコミュニティ 2 を生成する。

各コミュニティにおける全生成者の初期値は等しく差がないものとする。さらに、生成者が生成する生成物についても同様

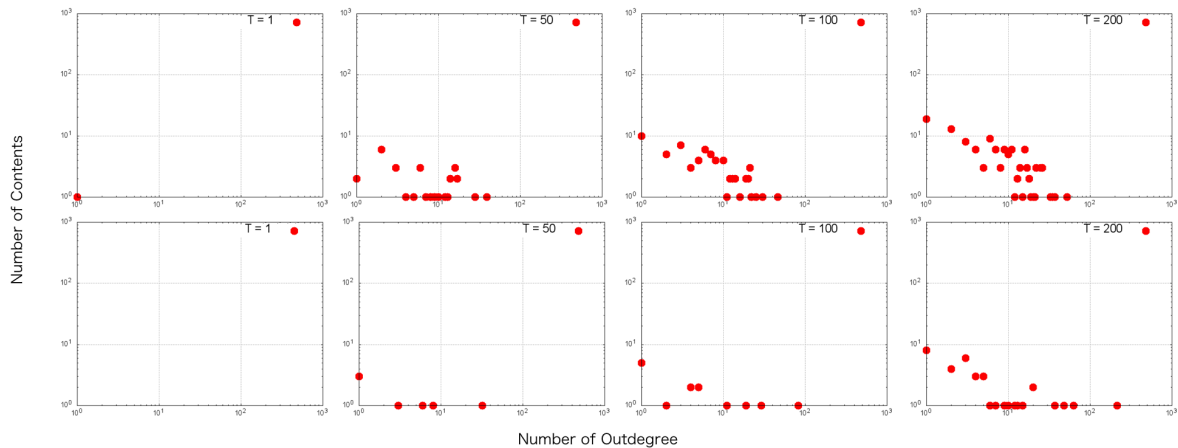


図 3 環境特性から見たウェブ空間成長

にネットワーク構造を観察するための outdegree, indegree 以外の特性は等しいものとする。

生成者の ICT リテラシーは, コミュニティの環境特性や個人の経験・学習に依存する。本実験では, ICT 教育以外の特性は等しいため生成者の ICT リテラシーと環境特性の ICT 教育の相互作用が生成者群による情報発信活動の活性度にもたらす影響を観察することができる。

図 3 は前述の設定を用いてシミュレーションを行なった結果を outdegree 図で表現したものである。上段にコミュニティ 1, 下段にコミュニティ 2 を表示し, 左から右にグラフを時系列順に表示してある。

コミュニティ 1: 初期状態 $T=1$: outdegree 図にコンテンツはほぼ確認できない。 $T=50$: 情報発信が活性化し outdegree 図で確認することができるコンテンツ数が増える。 $T=100$: さらに outdegree 図は発展し, 小さいがべき的減衰構造を確認することができる。 $T=200$: outdegree 図の構造に大きな変化はないがコンテンツ数が増加し outdegree 図の密度が増している。

コミュニティ 2: 初期状態 $T=1$: outdegree 図にコンテンツを確認することはできない。 $T=50-100$: 少数ではあるがリンクを持ったコンテンツを確認することが可能になるが, コンテンツ間の繋がりはまだ弱い。 $T=200$: outdegree 図が発展している様子がみられ, べき的減衰構造が少し確認できるようになる。

今回実施したシミュレーションでは, 各コミュニティごとに ICT 教育の質を変え, 生成者の ICT リテラシーの向上に差が出る状況を再現した。

コミュニティ 1 では, ICT 教育の質が良く生成者群の ICT リテラシーが早熟であった。そのため早期で outdegree 図の構造に変化が現れべき的減衰構造を確認することができた。

コミュニティ 2 では, ICT 教育の質が悪く生成者群の ICT リテラシーの向上に時間を要した。そのため, $T=200$ の段階でようやくべき的減衰構造を確認することが可能になった。

以上より, ICT 教育の質が良く, コミュニティ内生成者の ICT リテラシーが向上しやすい環境が整えてあるコミュニティでは, ICT 教育の質が悪いコミュニティに比べて生成者の情報発信活動が活発になり, コミュニティ内のコンテンツ数やコンテンツ間の繋がりが多くなるとことが確認された。

コミュニティの ICT リテラシー教育の質が生成者軍の ICT リテラシーの影響を与える場合, ICT リテラシー教育の良し悪しがコミュニティ内の情報発信活動を活性化しウェブ空間の成長にも影響を与えることがわかった。そのため, コミュニティの情報発信活動を活性化させるための手法として ICT リテラシー教育に投資を行い教育の質の向上を図ることが有効であることを示唆した。

6 まとめと今後の課題

本稿では, コミュニティの ICT 教育が生成者群の成長を促すと考え, ICT 教育の異なる 2 つのコミュニティを生成し, 両者の情報発信活動を観察した。今回の実験では, 小規模なコミュニティを想定し, そのコミュニティ内での情報発信活動を観察したが, 今後は, エージェント数を拡大し大規模なシミュレーションを行うことを目指す。さらに, 生成者の ICT 学習モデルの構築や, ICT リテラシーや ICT 教育以外のパラメータが及ぼす影響の調査を行っていく。

参考文献

- [1] 岩田 翔太, 中平 勝子, 北島 宗雄: 通信基盤と人との相互作用による情報生成・流通プロセスの記述とバイラルメディアへの適応, FIT2014(第 14 回情報科学技術フォーラム)p337 p338, 2014.
- [2] Albert, A., Jeong, H., and Barabasi, A.-L.: Diameter of the World Wide Web, Nature 401, p130 p131, 1999.
- [3] A.-L. Barabasi, R. Albert, H. Jeong, G. Bianconi: Power-Law Distribution of the World Wide Web, SCIENCE VOL 287 p481 p484, 2115a 2000.
- [4] 福田 健介: ウェブ構造はどこまでスケールフリー/スモールワールドか~ウェブ構造のモデル化に向けて~, 情報処理学会研究報告, 136 号, p17 p22 2004,
- [5] 熊澤 健人, 中平 勝子, 北島 宗雄: outdegree 図に着目した, ウェブリンク構造変容分析によるデパイド状態の評価, FIT2016(第 15 回情報科学技術フォーラム), 2016.
- [6] 石原 直幸, 中平 勝子, 三上 喜貴: ccTLD を単位とした Web コミュニティ構造の分析, FIT2007(第 6 回情報科学フォーラム), p115 p118, 2007.
- [7] 三上 喜貴, 中平 勝子, 児玉 茂昭: 言語天文台から見た世界の情報格差, 慶応大学出版会, 2014.