

# 計量テキスト分析を用いた「生命の定義」を構成する要素とその関係性に関する研究

原田 潤<sup>†</sup>

## A study of the components of definitions of life and their relationships using quantitative text analysis

Jun Harada

### 1. はじめに

300を超える生命の定義が提案されている[1]と言われているが、未だ合意されているものは無く[2]、定義するのは不可能である[3]とも考えられている。一方、環境倫理や生命倫理などの倫理的問題の解決[3]や人工生命 (ALife)、合成生物学、宇宙生物学、生命の起源に関する研究[4]などの分野では、生命の定義は必要であると考えられている。

生命の定義は不可能であるとしても、生命という概念がどのように捉えられているのかを理解することは、生物学をはじめとした生命という概念を扱う分野において重要であると考えられる。

Trifonovは、123の生命の定義を構成する単語を抽出し、それらの単語は10のグループに分類できることを示した[5]。単語分類を行い構成されたグループは、それぞれ LIFE, SYSTEM, MATTER, CHEMICAL, COMPLEXITY, REPRODUCTION, EVOLUTION, ENVIRONMENT, ENERGY, ABILITYと命名された。例えば、LIFEグループに分類される単語は、“life”を含め、“living”, “alive”, “being”などの“life”を含意する単語である。そして、これらの単語の関係性を考察した上で、最小限の単語による生命の定義として、“Life is self-reproduction with variations” (生命とは変異を伴う自己複製である) と提案した。生命の定義は、生物学を中心とした分野において提案されているため、Trifonovの単語分類グループは生物学における生命概念の要素が概ね反映されているものと考えられる。

また、Malaterre & Chartierは、生物学関係の論文を以下のような手順でテキストマイニングすることによって、生物学的対象の生命らしさ (lifeness) の度合を評価している[6]。1969年から2012年の間に出版された54誌30,622報の生物学関連の論文に対してテキストマイニングを行い、

それらに出現する単語を200のクラスターに分類し、各クラスターを生物学におけるトピックとした。さらにこれら200のトピックを11のカテゴリに分類した。各トピックと、生命らしさ (lifeness) の度合を評価したい生物学的対象 (細菌, 古細菌, ウイルス, フェージ, プラスミド, プリオン, アデニン分子) を表す単語の集合との関連性を定量化することで、11のカテゴリのうち8つのカテゴリ (①Metabolism, ②Catalysis and synthesis of catalysis, ③Elaboration of cellular / structural features, ④Micro / macro environment (interactions), ⑤Plants / animals related (interactions), ⑥Humans related (interactions), ⑦Evolvability, ⑧Information encoding and genetics) が生命らしさ (lifeness) に関連することを示した。さらに8つのカテゴリは、①MATTER-ENERGY, ②STRUCTURE, ③ENVIRONMENT INTERACTIONS, ④EVOLUTION, ⑤INFORMATION-PROGRAMの5つのSubject matterに分類された。これらのSubject matterは、Trifonovが分類した生命の定義に出現する単語のグループにも出現していることから、生命の定義には生物学における生命概念がある程度表現されていると考えられる。

以上の研究は、生命の定義や生物学論文を分析し、それらに出現する単語をグループやSubject matterといった項目に分類したものである。それらを生命概念の構成要素として捉えた場合、生命概念の構成要素間の関係性については十分な検討がなされていない。生命概念がどのように捉えられているか理解するためには、生命概念の構成要素間の関係性についても明らかにする必要がある。

本研究では、Trifonovが分類したグループやMalaterre & Chartierが示したSubject matterに出現するような項目を生命概念の構成要素と捉え、これまで提案されてきた生命の定義に対して計量テキスト分析による内容分析を行うことで生命概念を構成する要素とその関係性を検討した。計量テキスト分析には、計量テキスト分析のためのフリー

<sup>†</sup>2023年度修了 (自然環境科学プログラム)

## 計量テキスト分析を用いた「生命の定義」を構成する要素と その関係性に関する研究

ソフトウェアであるKH Coder[7]を用いた。計量テキスト分析の結果を基に生命の定義に出現する単語を抽出し、それらをTrifonovが行った単語分類の10のグループに分類して、グループ間の共起ネットワーク図の作成および階層的クラスタ分析を実施し、共起の強さやクラスタから関係性を調べた。

## 2. 研究対象および方法

### 2.1 研究対象とその収集方法

本研究の研究対象は、「生命の定義」である。Koh and Lingは、Trifonovが分析対象としたBarbieriおよびPopaの収集した生命の定義[8], [9]に加え、2011年までに発表された合計135の生命の定義をリスト化している[10]。本研究では、Koh and Ling, Barbieri, Popaのリストに加え、独自に文献データベースによって生命の定義を収集した。

文献データベースは、Springer Link, Science Direct, Cambridge Core, Pub Medを用いた。各データベースにて、“definition of life”, “defining life”, “what is life” というフレーズにて検索を行い、ヒットした文献のアブストラクトを確認し、生命の定義が記載されている場合は、その定義をリストに加えた。Springer Link, Science Direct, Pub Medでは、タイトルの検索が可能のため、タイトル検索を行い、Cambridge Coreについては全文検索を行った。また、上記の手順によって収集した定義以外にも、研究を実施するにあたって確認した論文や書籍の中で取り上げられている定義を収集した。

どのような記述が生命の定義であるかの判断については、次のように行った。前述のKoh and Ling, Barbieri, Popaが生命の定義としてリスト化しているものについてはそのまま使用し、独自に検索し収集したものについては、文献の中で“life is defined～”, “life is～”など明確に生命とは何かということを簡潔に記述している部分を抜粋した。主語の部分は、“life”の他、“living system”, “living entities”, “living things”などの生命に類似する表現であるものも含めている。

収集した生命の定義に関する記述には、定義ができないという内容や科学的な表現ではない記述もあり、このような記述を除外した生命の定義のリストを作成し、分析対象とした。分析対象の生命の定義の数は153であった。

収集した生命の定義は全て英語で記述されたものであるが、先行研究に挙げたTrifonovおよび、Malaterre & Chartierの結果から、生命の定義には生物学における生命概念の構成要素が概ね含まれていると考えられること、生物学の論文の約87%が英語で発表されている[11]ことから、本研究で収集した生命の定義には、生物学における生命概念が概ね表現されているであろうと考えられる。

### 2.2 計量テキスト分析の概要

生命の定義はテキストデータであり、その内容は生物学

の知見を基にそれぞれの研究者が考える生命概念が記述されていると考えられることから、生命の定義に対して内容分析を行うことで、生命の定義の構成要素やそれらの関係性を知ることができると考えられる。

以上より、収集した生命の定義に関する記述について、計量テキスト分析を行った。計量テキスト分析とは、樋口によると「計量的分析手法を用いてテキスト型データを整理または分析し、内容分析 (content analysis) を行う方法である。計量テキスト分析の実践においては、コンピュータの適切な利用が望ましい。」と述べられている[7]。

計量テキスト分析には、計量テキスト分析のためのフリーソフトウェアであるKH Coderを用いた。計量テキスト分析の結果を基に生命を定義に出現する単語とその出現頻度および、それらをTrifonovが行った単語分類に沿って10グループに分類した上で共起関係を調べた。

### 2.3 形態素解析の方法

生命の定義に使用されている単語を定量化し分析するために形態素解析を実施した。形態素とは言語において意味を持つ最小単位であり、形態素に分割し、品詞などを判別することを形態素解析という。例えば、“I am a student.”という英文は、“I / am / a / student / .”といったように分割できる。一般的に形態素解析では、対象の言語の辞書を用いて行われる。本研究の形態素解析は、KH Coderに内蔵されている英語の形態素解析エンジンであるStanford POS Taggerを使用した。

KH Coderにおける形態素解析では、stop wordsという抽出しない単語を設定することができる。一般的に英文の分析ではbe動詞のような文章にも出現する単語をstop wordsとして指定する。本研究の分析では、KH Coderに同梱されているstop wordsのサンプルから“being”および“system”の2語を除外し使用した。この2語を除外した理由としては、どちらも先行研究に挙げたTrifonovの単語分類において出現することから、生命の定義において意味を持つ単語であると考えられるためである。また、生命の定義内に出現する[] (角括弧) については、“-RSB”や“-LSB-”といった形で抽出される。同様の語として“-lsb\_NN”, “-rsb\_NN”, “-lsb\_VB”がある。これら5種類の語は、生命の定義の内容に直接関係はしないため、KH Coderの「分析に使用する語の取捨選択」コマンドで使用しない語として指定した。

### 2.4 コーディング方法

KH Coderでは、ある文書に特定の単語が出現していた時、その文書にコードを付与し分析することができる。例えば、分析対象とした生命の定義には、“self-reproducing”, “self-reproduction”, “reproduction”, “reproduce”という単語が出現している。これらの単語は、「自己複製」という共通の概念にまとめることができる。計量テキスト分析では、この概念をコンセプトと呼

び、このコンセプトを取り出す処理をコーディングと呼ぶ[12]。KH Coderでは、「指定した条件が満たされれば、あるコンセプトが出現していたと見なす」[12]というルール（コーディングルール）を設定することで、コンセプトを用いた分析が可能となる。上記を例にすると、“self-reproducing”, “self-reproduction”, “reproduction”, “reproduce” のいずれかの語が出現した場合、「自己複製」というコンセプトが出現していたと見なすコーディングルールを設定することで、それらの単語が出現している生命の定義に「自己複製」というコンセプトが出現していたと見なし分析することができる。本研究では、Trifonovによる単語分類の10のグループをコンセプトとしてコーディングルールを設定した。例えば、本研究で用いた生命の定義の一つであるJoyceによる定義“Life is a self-sustained chemical system capable of undergoing Darwinian evolution.” [13]には、LIFE (Life), SYSTEM (system), CHEMICAL (chemical), EVOLUTION (evolution), ABILITY (capable) の5つのコンセプトが出現する（カッコ内はそれぞれのコンセプトに該当する単語）。

本研究では、コーディングを行い付与したコンセプトを基に共起ネットワーク図の作成と階層的クラスタ分析を行った。

## 2.5 階層的クラスタ分析の方法

KH Coderにおいて、階層的クラスタ分析は「出現パターンの似通った語の組み合わせにはどんなものがあったのか」[7]を探索する手法である。出現パターンの似通った語、すなわち共起の程度が強い語同士をまとめてクラスタが形成される。本研究では、前節の方法でコーディングを行って付与したコンセプトを用いて階層的クラスタ分析を行った。それにより、コンセプト間の共起の程度に基づいてどのコンセプトがどのコンセプトと類似しているか知ることができる。クラスタ分析の方法はWard法、出現パターンの類似度（共起の程度）はJaccard係数を用いて分析を実施した。クラスタ数を決定するために併合水準（非類似度）を確認し（図1）、併合水準のグラフの傾きが急激に上昇する直前のクラスタ数である5のクラスタでデンドログラム（樹状図）を作成した。この傾きの急激な上昇は、非類似度が急激に上昇すること、つまり類似していないものも無理にまとめることを意味するため、その直前のクラスタ数である5が適していると考えられる。

## 3. 結果

### 3.1 形態素解析

153の生命の定義に対して行った形態素解析によって抽出された単語（抽出語）は、延べ数として2,553語であり、語の種類（異なる語）としては1,004種類であった。これら1,004種類の単語はstop wordsおよび、使用しない語として指定した5種類の単語を除き、形態素解析エンジ

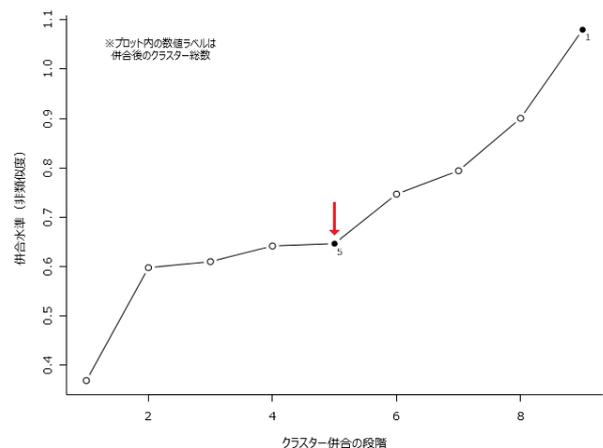


図1 クラスタ数決定のための併合水準（非類似度）

ンであるStanford POS Taggerにより互いに異なる単語と見なされた単語である。なお、同一単語・同一品詞の活用形は同一単語としてカウントされる。例えば、“system”, “systems” はどちらも“system”としてカウントされる。

比較的出現頻度、出現文書数（定義数）が多い出現頻度が10回以上および出現文書数が10文書以上の単語を表1と表2に示す。出現頻度の最頻値は1回、平均値は2.5回、標準偏差は5.4回であり、出現文書数の最頻値は1文書、平均値は2.3文書、標準偏差は4.3文書であった。

出現頻度上位3位の単語（“system”, “Life”, “life”）が特に出現頻度が高かった。「生命の定義」を分析対象としたため、“Life”, “life”, “living”, “live”, “alive”などの生命に関係する単語や、“define”, “definition”などの定義に関する単語が上位に出現していたと考えられる。

Trifonovが示した生命の定義に出現する単語を分類したグループ（LIFE, SYSTEM, MATTER, CHEMICAL, COMPLEXITY, REPRODUCTION, EVOLUTION, ENVIRONMENT, ENERGY, ABILITY）に属する単語は、いずれも出現頻度が10回以上および出現文書数が10文書以上の単語に出現していた。

### 3.2 コンセプトを用いた共起ネットワーク図の作成

Trifonovが分類した生命の定義に出現する単語のグループであるLIFE, SYSTEM, MATTER, CHEMICAL, COMPLEXITY, REPRODUCTION, EVOLUTION, ENVIRONMENT, ENERGY, ABILITYをコンセプトとして、コーディングルールを設定した。

Trifonovの論文には上記の10グループに属する単語のリストが掲載されている。このリストを用いて、それぞれのグループに属する単語のいずれかが出現した場合に、コンセプトとしてそのグループが出現したと見なすというコーディングルールを設定した。例えば、TrifonovはLIFEに分類される単語として、“life”, “living”, “alive”, “being”などを挙げていることから、KH Coderにて、“life”, “living”, “alive”, “being”・・・のいずれかの単語が出現した場合、LIFEというコンセプトが出現したと見なすと

計量テキスト分析を用いた「生命の定義」を構成する要素と  
その関係性に関する研究

表1 出現頻度が10回以上の単語

抽出語	出現頻度	抽出語	出現頻度	抽出語	出現頻度
system	85	ability	15	evolve	11
Life	82	complexity	15	existence	11
life	73	definition	15	open	11
living	36	evolution	15	organic	11
process	30	information	15	organization	11
organism	29	molecular	15	able	10
live	27	entity	14	cell	10
environment	26	level	14	network	10
matter	26	material	14	state	10
energy	23	reproduction	14	thing	10
define	22	alive	13		
chemical	19	form	12		
capable	17	reproduce	12		
property	17	structure	12		
metabolism	16	characteristic	11		

表2 出現文書数（定義数）が10文書以上の単語

抽出語	文書数	抽出語	文書数	抽出語	文書数
Life	72	evolution	14	form	10
system	57	information	14	material	10
life	53	ability	13	organization	10
living	29	alive	13		
live	26	definition	13		
matter	22	reproduction	13		
process	22	complexity	12		
define	20	entity	12		
organism	20	reproduce	12		
energy	19	structure	12		
chemical	17	characteristic	11		
environment	17	evolve	11		
capable	16	molecular	11		
property	16	open	11		
metabolism	15	existence	10		

いうコーディングルールを設定した。

なお、Trifonovの単語リストには含まれていないが、属すると思われる単語があった場合、コーディングルールに追加した。例えば、TrifonovはCHEMICALに属する単語の一つに“metabolism”を挙げているが、本研究で対象とした生命の定義の中には“metabolic”や“metabolically”といった同様の意味の単語が出現している。これらはTrifonovの単語リストには含まれていないため、これらの単語もCHEMICALのコーディングルールに追加した。

これらのコンセプトを用いて共起ネットワーク図を作成した(図2)。共起ネットワークとは、「共起の程度が強い語を線で結んだネットワーク」[7]である。共起の程度は、Jaccard係数を用いることで定量化できる。Jaccard係数は、「語Aと語B両方とも含む文書数/語Aまたは語Bを含む文書数」という式で求められる係数で、語Aと語Bの類似度の指標となる。Jaccard係数が大きいということは、2つの語の出現パターンが類似している、すなわち共起の程度が強いことを意味する。作成した共起ネットワーク図は、類似度が高い上位10位(Jaccard係数0.34以上)の共起ネットワーク図である。なお、KH Coderではお互い強く結びついている語同士はグループ化され、同じ色

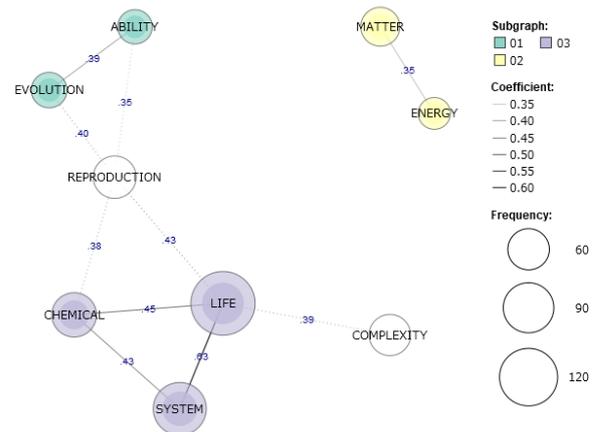


図2 コンセプトを用いた共起ネットワーク図  
(Jaccard 係数：0.34 以上)

で表現される。

LIFE, SYSTEM, CHEMICALが同一グループに共起していることから、生命は化学的なシステムであると定義されていることが伺える。また、LIFEとCOMPLEXITYが共起していることから、生命は複雑性を持つものとして定義されていることが伺える。ABILITYとEVOLUTIONが同一グループで共起していることから、進化に関する能力は生命を定義付ける要素の一つと考えられる。REPRODUCTIONはABILITY, EVOLUTION, LIFE, CHEMICALと共起していることから、自己複製は能力であり、進化に関わり、化学的な現象と関係することが伺える。MATTERとENERGYが同一グループに共起していることから、生命の定義において物質とエネルギーはひとまとまりで言及されていることが伺える。ENVIRONMENTは類似度の高い共起ネットワーク図に出現しなかったことから、環境関係の単語は特定のコンセプトに属する単語との共起関係が比較的弱いと考えられる。

### 3.3 コンセプトを用いた階層的クラスター分析

「2.5 階層的クラスター分析の方法」で述べたように併合水準（非類似度）を基にクラスター数を5と決定(図1)し、5つのクラスターを持つデンドログラム（樹状図）を作成した(図3)。四角い枠で囲まれたコンセプトが同一のクラスターに属する。

コンセプトを用いて階層的クラスター分析を行った結果、クラスターは、①ENVIRONMENT, ②MATTERとENERGY, ③ABILITYとREPRODUCTIONとEVOLUTION, ④COMPLEXITY, ⑤CHEMICALとLIFEとSYSTEM, となった。

この結果から、生命の定義において、生命は化学的なシステムであること、それは複雑性を有し、自己複製と進化という能力を持つこと、環境（物質とエネルギー）に関係すること、が伺えた。

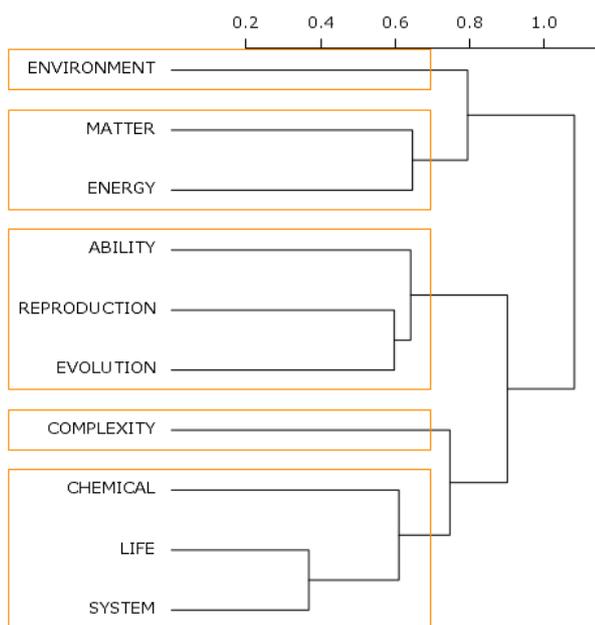


図3 コンセプトを用いた階層的クラスター分析のデンドログラム

#### 4. 考察

生命の定義に対して計量テキスト分析を実施した結果、“system”という単語の出現頻度が最も高いこと、“system”は出現定義数が“Life”に次いで多いことから、生命概念は多くの研究者にとって、ある種のシステムであると捉えられていることが伺えた。システムという概念自体も様々な定義が提案されており、共通の定義を構築することが課題とされている[14]。本研究で対象とした生命の定義で述べられているシステムは、“dynamic low entropy systems” [15] (動的な低エントロピーシステム)、“open non-equilibrium systems” [8] (開放型非平衡システム)、“self-sustained molecular system” [16] (自立した分子システム)、“non-equilibrium chemical system” [16] (非平衡化学システム)、“multilevel open catalytic system” [16] (多階層開放型触媒システム)、“ensemble of molecular-informational feedback-loop systems” [16] (分子-情報フィードバックループシステムの集合体)、“self-sustaining chemical system” [17] (自立した化学システム) など様々な表現がなされている。生命の定義において、生命はどのようなシステムであるかを示すことが焦点となっていると考えられる。コンセプトを用いた共起ネットワーク図 (図2) において、LIFE、SYSTEMに関係する単語とCHEMICALに関係する単語は共起関係が比較的に強いこと、同じくコンセプトを用いた階層的クラスター分析 (図3) においてLIFE、SYSTEM、CHEMICALは同一のクラスターであることから、生命の定義において生命は化学的なシステムであると捉えられていることが伺えた。

LIFEとCOMPLEXITYのJaccard係数は0.39であり比較的に類似度が高いこと、階層的クラスター分析では、COMPLEXITYは独立しているものの、CHEMICALと

LIFEとSYSTEMのクラスターに隣接していることから、生命の定義において、複雑性を有する化学的なシステムであることは生物の中心的な性質の一つと捉えられていることが伺えた。NASAに採用されているJoyceによる生命の定義では“Life is a self-sustained chemical system capable of undergoing Darwinian evolution.” [13] (生命とは、ダーウィンの進化を遂げることができる自立した化学システムである。)とあり、NASAにおいて生命は化学的なシステムと捉えられていることが伺える。また、複雑性に言及している例では、Gusevによる“Life is an energy-dependent chemical cyclic process which results in an increase of functional and structural complexity of living systems and their inhabited environment.” [16] (生命とは、エネルギーに依存した化学的な循環プロセスであり、その結果、生命システムとその生息環境は機能的、構造的に複雑さを増していく。)のような定義がある。

COMPLEXITYに分類される単語の中で最も出現頻度の多い単語は“information” (情報)であった。その他、COMPLEXITYに分類される単語には“code”, “program”, “signal”, “message”など情報関係の単語が多いことから、情報は生命の定義における複雑性の中心的概念であると考えられる。Dysonは、“Life is defined as a material system that can acquire store, process, and use information to organize its activities.” [9] (生命とは、その活動を組織化するために情報を獲得し、蓄積し、処理し、利用することができる物質的なシステムであると定義される。)と提案しており、情報の獲得・蓄積・処理・利用という観点で複雑性に関するシステムであるとされる。以上より、生命の定義において、生命は情報の獲得・蓄積・処理・利用のような複雑性を有する化学的なシステムであると捉えられていることが伺えた。

西田は情報学と生命論は本質的に表裏一体であることを論じた上で、情報と生命という概念自体の関係性も表裏一体であるという主張にもある程度の妥当性があるだろうと述べている[18]。また、西垣は、情報の本質は生物による意味作用であり、「情報 (information) とは生命体の外部に実体としてあるものではなく、刺激を受けた生命体の内部 (in) に形成 (form) されるものである。」と述べている[19]。以上のように、情報と生命は密接な関係があると考えられるが、生命と同様、情報も様々な定義が提案されている[20]。生命の定義における複雑性の位置づけを明らかにするためには、生命の定義において複雑性の中心的概念であると考えられる情報概念がどのような意味で使用されているかについて個別に分析する必要があると考えられる。

共起ネットワーク図 (図2) において、LIFEとREPRODUCTIONのJaccard係数は0.43と比較的に高い類似度を示しており、自己複製は生命の定義において主要な要素であることが伺える。また、REPRODUCTIONは、EVOLUTION、ABILITYとも比較的に高い類似度を示しており、階層的クラスター分析 (図3) でもABILITYと

計量テキスト分析を用いた「生命の定義」を構成する要素と  
その関係性に関する研究

REPRODUCTIONとEVOLUTIONは同じクラスターになっていることから、自己複製や進化は生命というシステムの有する能力であると捉えられていると考えられる。自己複製や進化が可能であるという点について記述している定義としては、Weberによる“Living entities can be viewed as bounded, informed autocatalytic cycles feeding off matter/energy gradients, exhibiting agency, capable of growth, reproduction, and evolution.” [21] (生命体は、物質/エネルギー勾配を利用し、成長、複製、進化が可能であり、主体性を発揮する、境界を持つ、情報に基づいた自己触媒サイクルとみなすことができる。) という定義がある。

共起ネットワーク図(図2)において、ENERGYとMATTERは共起の程度が比較的強く同一グループであること、階層的クラスター分析(図3)においてENERGYとMATTERは同じクラスターであることから、生命の定義において両者はひとまとまりに言及される概念であることが伺えた。生物は細胞から構成されており、細胞は水、タンパク質、脂質などの物質で構成されている。また、物質を代謝することでエネルギーを産生し生命活動が行われる。物質とエネルギーは生物の体を構成し、活動する上で不可欠であるため、生命の定義においてひとまとまりに言及される傾向にあると考えられる。Fongの“Life is made of three basic elements: matter, energy and information... Any element in life that is not matter and energy can be reduced to information.” [22] (生命は、物質、エネルギー、情報の3つの基本要素からできている...物質とエネルギー以外の生命の要素は、情報に還元することができる。) や Baltscheffskyの“Life may [...] be described as a flow of energy, matter and information.” [23] (生命は、エネルギー、物質、情報の流れであるとも言えるかもしれない。) といった定義にもあるように、エネルギーと物質は生命の基本要素として捉えられている。また、Varelaの定義“A physical system can be said to be living if it is able to transform external energy/matter into an internal process of self-maintenance and self-generation.” [8] (物理システムが生きていると言えるのは、外部のエネルギーや物質を、自己維持と自己生成の内部プロセスに変換することができる場合である。～) や Baltscheffskyの定義“Life is defined as a self-sustained molecular system transforming energy and matter, thus realizing its capacity of replication with mutations and anastrophic evolution.” [16] (生命とは、エネルギーと物質を変換することによって、突然変異とanastrophicな進化を伴う複製能力を実現する、自立した分子システムと定義される。) にあるように、生命はある種のシステムであるが、そのシステムには物質的な基盤があり、物質やエネルギーを利用し、代謝や自己複製、進化といった能力を発現すると考えられる。

階層的クラスター分析(図3)において、ENERGYとMATTERのクラスターは、ENVIRONMENTに隣接して

いた。Lahav and Nirは、“A living entity is an ensemble of molecules which exhibit spatial organization and molecular-informational feedback loops in utilization of materials and energy from the environment for its growth, reproduction and evolution.” [24] (生命体は、その成長、生殖、進化のために環境から物質とエネルギーを利用する空間的な組織化と分子-情報フィードバックループを示す分子の集合体である。) と定義していることや、Trifonovはエネルギー供給と物質供給は環境を意味する、と述べている[5]のように、エネルギーと物質は環境から取り入れたり、相互作用したりするものと考えられる。また、環境はある主体に対するその周囲とされ、その把握は生物の反応を通してなされる[25]とされる。つまり、環境は生物・生命システムにとっては外部を意味する。Vitas and Dobovišekは、“Life is a far from equilibrium self-maintaining chemical system capable of processing, transforming and accumulating information acquired from the environment.” [26] (生命は、環境から獲得した情報を処理し、変換し、蓄積することができる、平衡からかけ離れた自己維持する化学システムである。) と定義しており、環境という概念は生命の定義において直接的な生命の性質ではなく、付帯する概念として捉えられていると考えられる。

最後に、先行研究として挙げたTrifonovの結果との比較を行う。Trifonovは、生命の定義を構成する単語の分析で得られた結果から、“Life is metabolizing material informational system with ability of self-reproduction with changes (evolution), which requires energy and suitable environment.” (生命は、変化(進化)を伴う自己複製能力を持つ、代謝する物質的・情動的なシステムであり、エネルギーと適切な環境を要求する。) という定義が可能であること示した[5]。さらに、自己複製は代謝とシステムを意味し、代謝は環境(エネルギー、物質供給)を意味すること、複雑さ(情報)は単純なものから複雑なものへの進化の過程で変化(進化)を伴う自己複製の産物と考えることができることから、上記の定義は“Life is self-reproduction with variations.” (生命は変異を伴う自己複製である。) という表現に整理できると提案した[5]。

本研究で対象とした生命の定義は、Trifonovの研究で対象となった生命の定義も多く含んでおり、単語の分類も同研究を踏襲しているが、生命の定義の構成要素の関係性は一部異なった結果となった。これは、Trifonovの研究と本研究の大きな相違点である分析手法によるものと考えられる。Trifonovによる生命の定義の構成要素の関係性の考察は、計量的な分析が行われていないのに対し、本研究では計量テキスト分析という計量的な手法を用いて実施した。本研究で実施した計量テキスト分析は、同じデータを用いれば、誰が行っても同じ結果を得ることができ、「第三者の批判・検討・検証に耐えるオープンさを有しているという意味での客観性」[27]が高いと考えられる手法である。

本研究は計量テキスト分析を用いたことにより、比較的客観性の高い分析を行えたと考えられる。

## 5. 結論

153の生命の定義に対して計量テキスト分析を行った結果、生命概念は以下のような要素と関係性が示された。

- ① 情報の獲得・蓄積・処理・利用などの複雑性を有する化学的なシステムである。
- ② そのシステムは、自己複製、進化といった能力を有する。
- ③ 上記の性質や能力は、物質的な基盤があったり、エネルギーや物質を利用したりすることで発現する。
- ④ 物質、エネルギーは、そのシステムの外部である環境から取り込まれる。システムと環境は物質、エネルギーを介して相互作用する。

一方で本研究には次のような限界と課題があると考えられる。本研究は英語で記述された定義のみを対象にしているため、限定された範囲の定義を対象としている。また、本研究で示された生命の定義の要素はTrifonovの単語分類による10のグループを踏襲しているが、それらのグループに分類される単語であるシステム（SYSTEMに分類）や情報（COMPLEXITYに分類）といった概念自体も様々な定義や捉え方があり一義的に定まっていない。生命の定義の構成要素やその関係性についてより深く検討するには、Trifonovによる単語分類の妥当性の検証や、生命の定義の要素である個々の概念が生命の定義においてどのような意味や観点で使用されているのか分析を行う必要があると考えられる。

以上のような限界や課題があるものの、前述のように先行研究に挙げたTrifonovとMalaterre & Chartierの結果[5]、[6]から生命の定義には生物学における生命概念の要素が概ね含まれていると考えられること、Trifonovの分類は生物学における生命概念の要素が反映されているものと考えられること、生物学の論文の約87%が英語で発表されている[11]ことから、本研究で分析対象とした英語で記述された生命の定義においても、生物学における生命概念が概ね表現されているであろうと考えられる。

本研究では限定された範囲ではあるものの、Trifonovが対象とした生命の定義の数を超える数の生命の定義を収集し分析を行ったこと、計量テキスト分析という客観性の高い分析手法を用いたことにより、生命の定義の構成要素とその関係性についてある程度妥当性を持った結果が得られたのではないかと考える。

## 謝辞

本研究の遂行や本論文の作成にあたり、自然環境科学プログラムの指導教員として多大なご指導賜りました二河成男教授に深謝致します。同プログラムの加藤和弘教授には

本研究にあたりご助言を賜りました。ここに深謝の意を表します。また、本研究の遂行にあたり、交流の場を設けて頂き、ご助言頂いた同期の皆様に感謝の意を表します。最後に、家庭での研究時間を確保してくれた妻の協力無くして本論文は完成しませんでした。ここに感謝の意を表します。

## 文献

- [1] Love, Luca Gregory. Does the discovery of the mimivirus call into question attempts to define life? *Bioscience Horizons: The International Journal of Student Research*. 2018, vol. 11.
- [2] Cleland, Carol E., Chyba, Christopher F. Defining 'Life'. *Origins of life and evolution of the biosphere*. 2002, vol. 32, p. 387-393.
- [3] Machery, Edouard. Why I stopped worrying about the definition of life... and why you should as well. *Synthese*. 2012, vol. 185, no. 1, p. 145-164.
- [4] Diéguez, Antonio. Life as a Homeostatic Property Cluster. *Biological Theory*. 2013, vol. 7, no. 2, p. 180-186.
- [5] Trifonov, Edward N. Vocabulary of Definitions of Life Suggests a Definition. *Journal of Biomolecular Structure and Dynamics*. 2011, vol. 29, no. 2, p. 259-266.
- [6] Malaterre, Christophe, Chartier, Jean-François. Beyond categorical definitions of life: a data-driven approach to assessing liveness. *Synthese*. 2021, vol. 198, no. 5, p. 4543-4572.
- [7] 樋口耕一. 社会調査のための計量テキスト分析：内容分析の継承と発展を目指して：KH Coder OFFICIAL BOOK. 第2版, 京都, ナカニシヤ出版, 2020, ISBN978-4-7795-1474-6.
- [8] Barbieri, Marcello. *The Organic Codes: An Introduction to Semantic Biology*. Cambridge University Press, 2002, ISBN978-0-511-61401-9.
- [9] Popa, Radu. *Between Necessity and Probability: Searching for the Definition and Origin of Life*. Springer-Verlag Berlin Heidelberg, 2004, 252p., ISBN978-3-642-05808-0.
- [10] Koh, Yong Zher, Ling, Maurice HT. On the Liveliness of Artificial Life. *Human Level Intelligence*. 2013, vol. 3, p. 1:1-17.
- [11] Monge-Nájera, Julian, Nielsen, Vanessa. The countries and languages that dominate biological research at the beginning of the 21st century. *Revista De Biología Tropical*. 2005, vol. 53, no. 1-2, p. 283-294.
- [12] 樋口耕一, 中村康則, 周景竜. 動かして学ぶ! はじめて

計量テキスト分析を用いた「生命の定義」を構成する要素と  
その関係性に関する研究

- のテキストマイニング: フリー・ソフトウェアを用いた自由記述の計量テキスト分析. 京都, ナカニシヤ出版, 2022, ISBN978-4-7795-1639-9.
- [13] Joyce, Gerald F. "Foreword" . *Origins of life : the central concepts*. Deamer, David W., Fleischaker, Gail R.編. Boston, Jones and Bartlett Publishers, 1994.
- [14] Sillitto, Hillary, Dori, Dov, Griego, Regina M., Jackson, Scott, Krob, Daniel, Godfrey, Patrick, Arnold, Eileen, Martin, James, McKinney, Dorothy. *Defining "System" : a Comprehensive Approach*. INCOSE International Symposium. 2017, vol. 27, no. 1, p. 170–186.
- [15] Prigogine, I. *From being to becoming : time and complexity in the physical sciences*. San Francisco, W. H. Freeman, 1980, ISBN978-0-7167-1107-0.
- [16] "Short definition of Life" . *Fundamentals of life*. Pályi, Gyula, Zucchi, Claudia, Caglioti, Luciano編. Paris, Elsevier, 2002, p. 15–55, ISBN978-2-84299-303-0.
- [17] Lazcano, Antonio. *Towards a Definition of Life : The Impossible Quest?* *Space Science Reviews*. 2008, vol. 135, no. 1–4, p. 5–10.
- [18] 西田洋平. 情報学の哲学的前提と生命観—メタ理論としての情報学と生命論の表裏一体性—. *情報メディア研究*. 2011, vol. 10, no. 1, p. 63–74.
- [19] 西垣通. *基礎情報学：生命から社会へ*. 東京, NTT出版, 2004, 235p., ISBN978-4-7571-0120-3.
- [20] 杉野隆. "『情報』の新たな定義を考える" . *情報システム学会 全国大会論文集*. 一般社団法人 情報システム学会, 2016.
- [21] Weber, Bruce H. *What is Life? Defining Life in the Context of Emergent Complexity*. *Origins of Life and Evolution of Biospheres*. 2010, vol. 40, no. 2, p. 221–229.
- [22] Fong, P. "Thermodynamic and Statistical Theory of Life: An Outline" . *Biogenesis Evolution Homeostasis*. Locker, A.編. Berlin, Heidelberg, Springer Berlin Heidelberg, 1973, p. 93–106, ISBN978-3-540-06134-2.
- [23] Baltscheffsky, Herrick. Major "Anastrophes" in the Origin and Early Evolution of Biological Energy Conversion. *Journal of Theoretical Biology*. 1997, vol. 187, no. 4, p. 495–501.
- [24] Lahav, Noam, Nir, Shlomo. "Life' s definition : in search for the most fundamental common denominators between all living entities through the entire history of life" . *Fundamentals of life*. Pályi, Gyula, Zucchi, Claudia, Caglioti, Luciano編. Paris, Elsevier, 2002, p. 131–133, ISBN978-2-84299-303-0.
- [25] 巖佐庸, 倉谷滋, 斎藤成也, 塚谷裕一. "環境" . *岩波生物学辞典*. 第5版, 東京, 岩波書店, 2013.
- [26] Vitas, Marko, Dobovišek, Andrej. *Towards a General Definition of Life. Origins of Life and Evolution of Biospheres*. 2019, vol. 49, no. 1–2, p. 77–88.
- [27] 樋口耕一. 計量テキスト分析およびKH Coderの利用状況と展望. *社会学評論*. 2017, vol. 68, no. 3, p. 334–350.