

自然言語処理による情報検索を用いた故障発想支援の提案

A Method to Support Failure Mode Identification by Using Information

Retrieval with Natural Language Processing

国立研究開発法人 宇宙航空研究開発機構 研究開発部門 第三研究ユニット

Japan Aerospace Exploration Agency, Research and Development Directorate, Research Unit III

○波平 晃佑 梅田 浩貴¹⁾ 大久保 梨思子¹⁾ 植田 泰士¹⁾ 片平 真史¹⁾ 森崎 修司²⁾ 天笠 俊之³⁾○Kohsuke Namihira, Hiroki Umeda¹⁾ Naoko Okubo¹⁾ Yasushi Ueda¹⁾ Masafumi Katahira¹⁾ Shuji Morisaki²⁾ Toshiyuki Amagasa³⁾

Abstract In FMEA, issue reports and past analysis results, which are past knowledge, are significant for identifying specific failure modes that capture the system's characteristics. However, failure modes and the reasons why they were identified, which JAXA has accumulated in Goal Structuring Notation (GSN), have not been great leverage as past knowledge. We have some problems occurred when applying the information retrieval of stored knowledge using natural language processing; 1) It is not an effective way to provide past failure information and failure mode as it is 2) It is not easy to create train data for accuracy improvement. In this paper, we propose a method for information retrieval, train data creation, and reflecting the retrieval result for identifying failure mode.

1. はじめに

FMEA（故障モードと影響解析）は、システムにおいて発生する故障を識別し、その影響の致命度に基づく相対的な評価等を通して、故障に対する対策を検討する手法である^[1]。JAXA では、FMEAを行うことでミッションの達成に影響がある故障モードを識別している。故障モードを識別するためには、システムやソフトウェア（以下、SW という）の故障発想の支援が必要である。その方法には、失敗経験等の過去知見の活用、ガイドワードの活用、分析アイテムの具体化の3つと言われている^[2]。特に失敗経験として過去の不具合情報は、事象や原因、その発生メカニズムが適切に記録されている場合、システムの特性を捉えた固有の故障モードを作成することに有効である^[3]。

一方、JAXA では、SW やシステムの論理に関する FMEA として、上記の3つの支援をもとに導出した故障モードとその導出経緯を Goal Structuring Notation^[4]（以下、GSN という）形式で蓄積しているが、蓄積された GSN 形式の成果物を知見として十分に活用されていないという問題がある。製品固有の故障モードを抽出するためには、分析者が過去の情報を解釈し、自身の製品の問題に落とし込むことが必要となるが、過去の情報は、類似製品開発を経験した分析者の経験や知見を有する者が参照することが前提となった情報であることが多く、過去知見の参照は極めて属人的となるためである。

国立研究開発法人 宇宙航空研究開発機構 研究開発部門 第三研究ユニット

Research Unit III, Research and Development Directorate, Japan Aerospace Exploration Agency

茨城県つくば市千現 2-1-1 Tel: 050-3362-2805 e-mail:namihira.kohsuke@jaxa.jp

2-1-1 Sengen, Tsukuba, Ibaraki, Japan

1) 国立研究開発法人 宇宙航空研究開発機構 研究開発部門 第三研究ユニット

Research Unit III, Research and Development Directorate, Japan Aerospace Exploration Agency

2) 名古屋大学 大学院情報学研究科

Graduate School of Informatics, Nagoya University

3) 筑波大学 計算科学研究センター

Center for Computational Science, University of Tsukuba

【キーワード：】 FMEA、GSN、故障モード、自然言語処理、不具合検索

そこで JAXA では、過去 FMEA 活動を通して蓄積された過去の知見（故障モードとその故障モード導出時に着目した仕様情報）の検索を通して、故障発想を支援する方法を検討した。本稿では、情報検索を使った故障発想への支援を目的とした、情報検索方法や学習データ作成方法、検索結果の反映方法を提案する。

2. 提案手法の適用先業務：故障モード導出

提案手法の適用先の分析業務である、宇宙機システム SW に対する故障モード導出について概説する。一般に SW に対する FMEA では、FMEA を適用する単位（以下、アイテムという）を「機能」として故障を分析することが多いが、「機能」は抽象概念であり、また本来正しく実装されるはずであり故障しない（論理単体では故障しない）ため、故障モードを発想する難易度が高いとされる^[5]。加えて、宇宙機システムの場合は、機能に対して多数の役割を持ったコンポーネントで構成されることも多く、「機能」の故障を発想する際にはコンポーネントの連携を考慮しなければならない。つまり、入出力のインターフェースに関する共通的な故障（来る/来ない等の入出力値異常）に加え、コンポーネント間の連携方法の差異によって異なる故障を想定しなければならない。例えば、単一方向のデータ送信のコンポーネント連携では入力データの異常を想定するが、自動復旧機能のような他コンポーネント状態を監視し復旧指示するコンポーネント連携ではそれぞれのコンポーネントの状態を考慮した故障を想定しなければならない（誤検知/未検知等）。さらに、サイバーフィジカルシステムである宇宙機システムは外部環境と相互作用を持つため、「機能」の故障にはその外部環境からの影響や故障原因を考慮しなければならない（姿勢決定機能における恒星の見え方等）。よって、宇宙機システム SW に対する故障モード発想には、従来の「機能を構成する SW 処理」に加え、「外部環境の状態」や「システム/機能を構成するコンポーネント間の連携」を考慮する必要がある。

3. 過去知見の活用ニーズ

宇宙機システム SW での故障モードを発想する際に、分析者は関連する開発文書等を参照し「機能」の故障モードを発想することになる。このときその参照する情報や範囲は膨大となることが問題となる。そこで、過去の不具合情報や故障モード等の過去知見の活用が有効となる。それらの過去知見には、実際に発生した又は想定した故障に対して、それに至った発生条件や状態、入出力値に関する情報が含まれている。特に、単体物理故障や経年劣化を原因としたハードウェア故障と異なり、SW 故障発想には故障モード（故障状態）だけでなく故障に繋がる原因や条件を含む情報は有用な情報となる。分析者は、分析するアイテムと過去知見に含まれる条件の共通点を参考に故障モードを導出する。よって、分析者は参照情報が膨大であり又は一部明記されていないために従来見逃していた仕様情報に対して効率的に条件を探索することができるため、過去知見の活用にはニーズがある。

ここで重要な点として、JAXA で開発される宇宙機システムは基本的に 1 点物の生産であるため、量産製品と比較して潜在する故障発生条件は未知の場合が多いことを指摘しておく。それでも、過去の不具合や分析結果は条件が部分的に一致する部分があり有用な情報となり得る可能性がある。なぜならば、利用状況や環境が新規であり基本構成が同じ場合（探査衛星等）や、同一又は類似のミッション目的に対してシステム/SW が新規である場合（シリーズ観測衛星/ロケット等）といった部分的な新規/差分があるためである。宇宙機システムの故障発想においても、過去の不具合情報や想定した故障モードを十分に踏まえることは故障モード発想促進や識別漏れ防止に繋がるため、過去知見の活用ニーズがある。

4. 過去知見の情報検索の課題

JAXA では新規性が高い製品に対する FMEA において、蓄積された過去知見の検索を通して、故障発想を支援する方法を検討した。特に情報検索の技術として、自然言語処理に機械学習を適用した情報検索を検討した。今回対象とする不具合情報や故障モードは自然言語で表現されているためである。またキーワード完全一致による情報検索では、情報に含まれる語彙が偏る状況下で

は検索結果が全く得られない又は膨大となってしまう場合があり、近年では機械学習を用いた不具合情報の情報検索が有効であることが示されている^[6]。

しかし、自然言語処理を用いた情報検索の適用にあたり、以下の課題が発生した。

■課題1：不具合情報や故障モードをそのまま提示しても効果が低い

故障モード発想支援を目的とした情報検索は、故障モード追加に繋がる情報の提示が求められる。しかし、個別製品での具体的な事例として記載される不具合情報は、同義語等の表記ゆれや語彙省略も含まれるため、不具合情報をそのまま提示した場合、検索ユーザ側で自身の製品の利用状況やSW処理との関連性の推定しなければならない（一致点や相違点の分析等）。また、過去のGSN記法の故障モード導出の分析結果（5章参照）についても、システムからSWまでの多くの仕様情報や分析者の思考情報（例：取捨選択した理由/ガイドワード等）を含むため、検索ユーザはそこから必要な情報を読み解く負荷が大きい。

■課題2：検索精度の向上に必要な学習データの作成が困難である

自然言語処理に機械学習を用いた場合の特性として、検索精度を向上させるためには学習データ、特にドメイン固有表現やそれらの関係性に関する知識を大量に確保する必要があるが、文章を正規化や構造化する必要があるその学習データを生成する負荷は高い^[7]。それに加え、宇宙機システム開発では、不具合情報や開発文書の表現は同義表現や、開発プロセス工程に伴う抽象や具体表現が多数含まれているため、情報検索における語彙の関係推定が難しいという課題がある。

5. JAXAに蓄積された故障モード導出に関する過去知見の構造

提案手法の情報検索では、源泉データとしてJAXAに蓄積された故障モード導出に関する分析結果（以下、故障モード導出経緯という）を扱う。そのデータ構造について説明する。

先述した通り、宇宙機システムにおけるSWの故障を発想する際には、「外部環境の状態（上位仕様相当）」、「システム/機能を構成するコンポーネント間の連携（基本設計仕様相当）」、「機能を構成するSW処理（詳細設計仕様相当）」の考慮が必要となる。JAXAでは、故障モード導出の際、それらの仕様情報を識別しながら故障モードを導出し、それらの分析過程をGSN記法で蓄積している^[8]。そのGSN記法では、Contextには上位から下位への順に「外部環境の状態や条件」、「コンポーネント間の連携」、「SW処理」に関する仕様が表現されている（図2参照）。なお、宇宙機システムは高信頼性が求められるため、識別される仕様情報は従来からの違いや複雑さがある条件であり、それらは潜在リスクが高い可能性である仕様情報といえる。末端Goalには故障モードが表現される。よって、末端Goalの故障モードも含む最上位から最下層Context組み合わせは、対象システムで想定された故障とその影響が1つの故障シナリオとして表現されており、故障発生時の外部環境、コンポーネント、SW処理での各条件や状態の組み合わせが表現されている。

6. 提案手法

6.1. 提案手法の工夫点

先述の課題を踏まえて、自然言語処理を用いた故障モードの検索の仕組みを提案する。この提案手法は以下の工夫点を持つ。

■工夫点1：故障発想に必要な情報である「システムやコンポーネント連携、SW処理の条件」を検索ユーザへ提示する（課題1改善）

新規性の高い製品におけるSWの故障モードを導出する際に利用する検索として、SW処理や実装だけではなく、外部環境やコンポーネント間連携等のSW仕様の前提がどこまで一致しているかに着目した検索、つまり、GSN記法として最上位から末端までのGoalにあるContextに着目した検索の仕組みにした。

- 提示されたContextが一致している場合、その対象箇所は既に検討済みであるため、分析者が考えるべき範囲は他の箇所であり、新規に考えるべき箇所が限定されるメリットがある。
- 提示されたContextが一部一致していない箇所がある場合、その一致していない理由が過去製品と分析対象製品の相違点となるため、新規性を捉えた故障モードが作成される。
- 提示されたContextが完全に不一致の場合、分析対象製品にありえない故障シナリオとなる

ため除外する必要がある。

つまり、いずれの場合でも、分析者は提示された情報を元に新規の故障発想が促進される。過去知見として蓄積された Context の組み合わせを 1 つの情報として活用することにより、分析対象との利用状況やコンポーネント連携、SW 処理の差分である部分的な新規性に気づくことができ、見逃していた条件や故障モードの導出への貢献に期待できる。また、検索結果は過去に想定された 1 つの故障シナリオを示しており、分析者は分析対象システムの特性に合わせた類似の故障シナリオを分析していることになる。

■工夫点 2: ノード間関係性から学習データを作成する自然言語抽出パターンを定義した(課題 2 改善)

先述したとおり、JAXA での過去の故障モードの分析結果は、従来の GSN 記法に加え、故障モード導出の目的を考慮し構造化されている(5 章参照)。そのため、GSN 内にあるノードの位置やノード間の関係から、故障モード発想における各情報の位置づけと関係性を読み取れる。そのノード間関係から故障モード発想支援に必要な自然言語情報の単位(学習データ)を自動作成する。この学習データは自然言語処理におけるモデルへの言語データや検索文書として利用される。未知語/抽象的な語彙があった場合でも、周辺ノード情報を含む形で自然言語情報を複数作成できるため、語彙の関連性の特定をする学習データを増やすことができ、情報検索の精度を向上できる。

6.2. 提案手法の実施方法

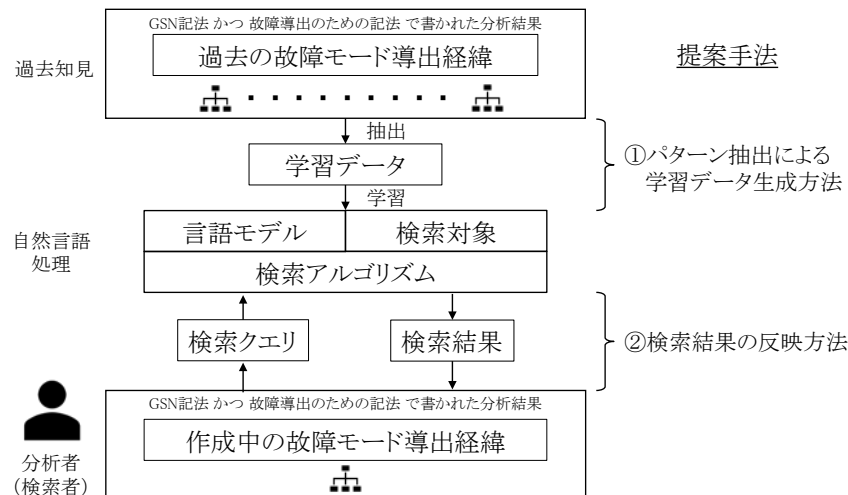


図1 提案手法の概要図

提案手法の概要を図1に示し、事前準備と検索利用時についてそれぞれ説明する。

①事前準備：自然言語処理/情報検索のための学習データの作成と学習方法

まず、提案手法は、Context 組み合わせに基づいて故障モード導出経緯内の自然言語を学習データ(テキストデータ)として抽出する(図2)。抽出された学習データを言語モデルの学習データや検索対象として利用する。今回、自然言語処理モデルとして Word2Vec^[9]を使い、検索アルゴリズムを cos 類似度による順位付けによる検索を利用した。つまり、抽出された学習データを使い Word2Vec (Skip-gram) モデルを事前に作成し、検索時には検索クエリと検索対象の cos 類似度を計算し順位付けを行う。そして、上位の順位の検索対象を検索ユーザに検索結果として提示する。

②検索と反映：検索結果を使った故障モード識別への反映方法

検索ユーザである分析者は、GSN 記法を使い、対象システムや SW の条件を識別しながら最終的に故障モードを識別することを目的としている。まず、検索ユーザは、検索結果として提示された Context (外部環境や SW 処理条件等) に対して、それまでに分析者が識別していた Context とを比較し{一致、部分一致、不一致}の条件を識別する(図3)。検索結果との差分に着目することにより、それまで見逃していた条件を Context として追加する。その差分は、分析対象における

検索結果参照の影響を軽減するため、情報量の少ない順である検索条件 2-1 から検索条件 2-3 の順で実施した。なお、検索クエリは、ユーザによる指定方法に依存しないように、仕様書上の機能名とした。例えば、“イベント検知機能”等である。

表 2 評価実施条件一覧

評価実施条件 No.	分析対象 (機能)	被験者 (検索ユーザ)		検索実施順序
1	機能 A	対象システム開発組織	2 名	1. 不具合情報 2. 故障モード
2	機能 A	第三者検証組織	2 名	1. 故障モード 2. 不具合情報
3	機能 B	対象システム開発組織	1 名	1. 不具合 2. 故障モード
4	機能 C	第三者検証組織	1 名	1. 故障モード 2. 不具合

表 3 各評価にて実施する検索条件一覧

検索条件 No.	不具合情報	故障モード		
	1	2-1	2-2	2-3
学習データに含まれる情報	外部環境から SW 処理まで含む 1 故障シナリオ	過去の故障モード	過去の故障モード導出時に考慮された SW 処理	外部環境から SW 処理まで含む 1 故障シナリオ
検索ユーザへの提示情報 (上位 10 件提示)	不具合情報	故障モード	Context + 故障モード	複数 Context + 故障モード
学習データ作成方法	不具合情報をそのまま提示	GSN からパターン抽出	GSN からパターン抽出	GSN からパターン抽出
作成されたデータ数 (検索対象)	103 件	328 件	537 件	316 件

7.2 結果と考察

■評価指標 1：影響大（ミッション継続不可）に繋がる故障モード数は増えるか？

従来の「不具合情報」「故障モード」や「SW 処理条件+故障モード」の検索に比べ、影響大に繋がる故障モードが多く追加されたため、「複数の Context を含めた故障モード」を用いた検索が最も効果のある検索方法と言える（表 4、図 4）。

また、検索結果を参照することで、従来見逃していたどのような条件が追加されたかの比率を図 5 にて示す。図 5 は、検索結果を参照し、「外部環境」「コンポーネント連携」「SW 処理」のどの種別の条件を検索ユーザが Context として新しく追加したかを計測した結果である。Context + 故障モードの場合に SW 処理条件および関連コンポーネントのに関する条件が追加された。システムから SW の条件を包含している複数 Context を提示した場合には、従来の SW 処理に関する条件だけでなく、外部環境等に関する条件も新しく追加された。また、不具合情報においてもシステムから SW の条件が含まれているが、SW の故障モード発想する分析者は SW 処理の条件のみに着目し外部環境に関する条件が追加されることはなかった。この場合、例えば、行き先の異なる探査衛星システムでの故障モード導出時といった、同一システム/アーキテクチャ/SW 処理において異なる外部環境に起因する故障モードを見逃す可能性がある。外部環境に関する条件に関する条件

が識別されない理由として、宇宙機システム SW の故障発想に必要な「外部環境」「コンポーネント連携」「SW 処理」の条件が必要であるが、今回の不具合情報は階層構造が不明確な「事象」「原因」「発生メカニズム」等の異なる視点で書かれており、検索ユーザは不具合情報から故障発想に必要な3つの条件を識別することが容易ではなかった。一方、提案手法である「複数 Context+故障モード」の検索手法では、情報がシステムから SW までの位置づけが明確に構造化されているため、条件識別が容易になったと言える。

表 4 識別された影響大に繋がる故障モードの数[個]。()内数字は識別された故障モードの総数。

評価実施 条件 No.	検索なし	検索あり：検索条件			
		1	2-1	2-2	2-3
	事前作成	不具合情報	故障モード	Context+ 故障モード	複数 Context+ 故障モード
1	3 (7)	7 (12)	4 (10)	5 (11)	7 (18)
2	5 (5)	5 (5)	5 (5)	6 (6)	7 (7)
3	2 (3)	6 (8)	4 (8)	6 (9)	9 (10)
4	4 (11)	4 (16)	5 (12)	7 (15)	9 (16)

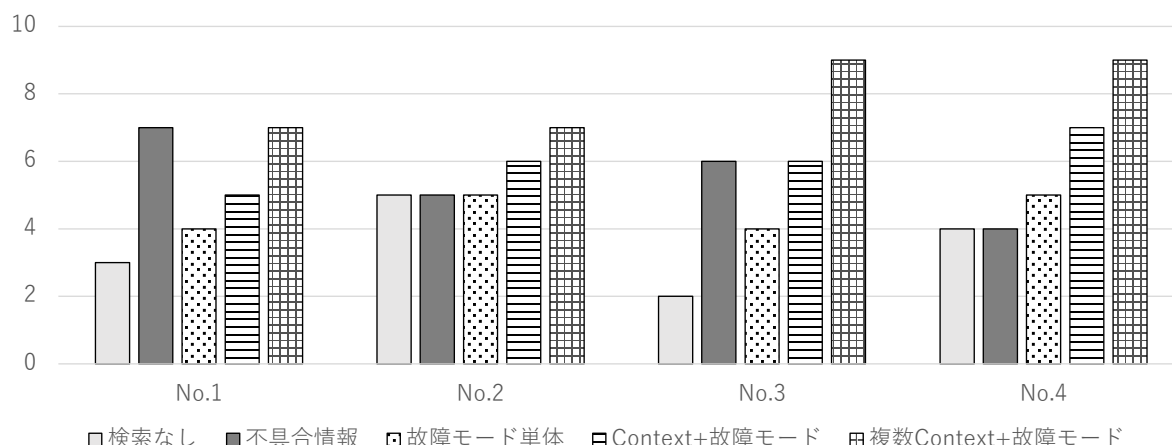


図 4 識別された影響大に繋がる故障モードの数[個]

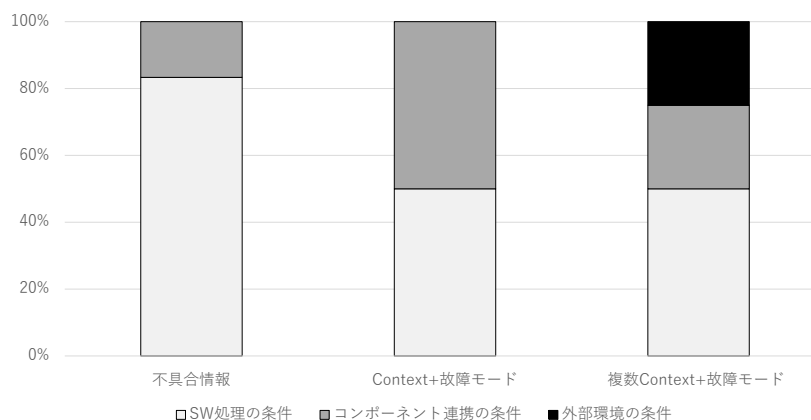


図 5 検索条件と追加された Context の位置づけの比率比較

■評価指標 2：検索ユーザの情報読み取り負荷は軽減するか？

本実験において、検索結果として提示される各情報に対して故障モードの導出に役に立つ箇所を特定するまでの時間を計測した。平均読み取り時間の結果を表 5 に示す。提案手法の検索では検索結果内に含まれる情報が増え、不具合情報に比べ情報量が多い。しかし、従来の不具合情報の読み取り時間で同等であるため、提案手法が提示する情報は読み取り負荷は低いといえる。また、評価指標 1 に示したように、提案手法は分析者が着目すべき仕様情報に対する気づきや故障モード導出に繋がっており、情報量が多くとも故障モード発想に有効な情報で構成されているといえる。

表 5 1 つの検索結果について役に立つ箇所を特定するまでの平均時間[分]

	不具合情報	故障モード	Context+ 故障モード	複数 Context+ 故障モード
平均文字数[文字数]	300	30	130	350
平均読み取り時間[分]	3.5	2.5	2.2	3.5

8. まとめ

GSN 記法かつ故障モード導出のためのモデリングルールをもつ過去の故障モード分析結果を起点とした検索により、故障モード導出に貢献できる仕組みを構築した。その検索手法は、故障モード発想に必要な「外部環境」「コンポーネント連携」「SW 処理」の条件の組み合わせに基づく学習データ生成方法および検索方法である。提案手法は、従来の不具合検索や故障モードのみの検索方法よりも、影響大の故障モードの導出が促進した。今後の展開として以下が考えられる。

- ・今回検索クエリは固定文字列としていたが、検索ユーザが作成中の GSN の情報をもとにクエリを補完する等の検索クエリを工夫することで、関連性の限定化やキーワードが思いつかない場合でも検索できるようにする等の検索精度向上を検討する。

- ・検索ユーザまたは業務結果からの検索結果へのフィードバック（学習データへのフィードバック）を取り込むことで、PDCA サイクルまで含めた検索手法および運用方法を検討する。

9. 参考文献

- [1] IEC 60812 Technical Committee: IEC 60812, analysis techniques for system reliability – procedure for failure mode and effects analysis (FMEA), 2018
- [2] 益田昭彦、新 FMEA 技法、河北印刷株式会社
- [3] 余宮尚志 他、不具合リスク発想のための観点の抽出方法とその効果、ソフトウェア品質シンポジウム 2016、2016
- [4] GSN COMMUNITY STANDARD VERSION 1 (<http://www.goalstructuringnotation.info/>)
- [5] H. Pentti, et al., Failure mode and effects analysis of software-based automation systems, in Proc. VTT Ind. Syst., 2002
- [6] P. Runeson, et al., Detection of Duplicate Defect Reports Using Natural Language Processing, 29th International Conference on Software Engineering (ICSE07), 2007
- [7] 土田正明 他、辞書とタグ無しコーパスを用いた固有表現抽出器の学習法、第 23 回人工知能学会全国大会、2009
- [8] 梅田浩貴 他、GSN 及び ESD モデルを用いたソフトウェア FMEA の提案、ソフトウェア品質シンポジウム 2018、2018
- [9] Mikolov Tomas, et al., Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013
- [10] IEC 61882: Hazard and operability studies (HAZOP studies)