Safety-IとSafety-IIからみた製品安全

早稲田大学理工学術院
創造理工学部 経営システム工学科 教授
小松原 明哲

最近、産業安全の世界では、安全に対して、Safety-I、Safety-IIという二つのアプローチが強く意識されるようになってきた。前者は信頼性に基づく安全の考え方である。後者は自然や社会環境下において、人の柔軟な行動により安全を求める考え方である。消費生活用製品などにおいての今までの製品安全は、多くは前者の考え方に基づいていたが、今後、増加が見込まれるAI搭載機器においては、後者の考え方に立った製品安全が求められる。

1. はじめに

最近、産業安全の世界では、安全に対して、Safety-I Safety-IIという二つのアプローチが強く意識されるようになってきた。Safety-I、Safety-IIは、ヒューマンファクターズの国際的権威であるE.Hollnagelが提唱した概念である[1]。前者は信頼性(リスク論)に基づく安全へのアプローチであり、リスクを低減することに主眼が置かれる。後者は状況に対する対応能力を向上することで安全と生産を首尾よく成就することに主眼が置かれ、レジリエンス・エンジニアリングを方法論とする。

家電製品等の一般消費生活用製品の使用に関わる製品安全については、今までもっぱらSafety-Iにより取り組まれてきたと思われる。しかし、今後、さまざまな形で展開が予想されるAI(人工知能)の関わる製品においては、むしろSafety-IIのアプローチが適合するように思われる。

本稿では、Safety-I、Safety-IIの概念を簡単に説明する。その上で、Safety-IIとDX技術との関係、製品安全においての課題について展望する。

2. Safety-I とSafety-II

2.1 Safety-I

(1) Safety-Iの考え方

Safety-Iとは端的に言えば、機械の取り扱いに関わる安全へのアプローチである。

機械はそれを設計した設計者がおり、その設計者が意図した通りに機械が機能する限りにおいて安全である。従って、機械それ自体が故障してはならず、かつ、オペレータが操作を誤ってはいけない。そこでこうした不具合(故障や誤操作)を減らす努力が求められる。具体的に言えば、機械の構成要素(部品)の信頼性を高め、安全上、重要な箇所は冗長化する等の対策を講じる。人間側についても同様で、教育訓練などによりオペレータの信頼性を高め、安全上重要な箇所は、フールプルーフやダブルチェックなどの対策を講じる。これらにより、機械(システム)全体の信頼性＝1を目指す活動が展開される(図1)。

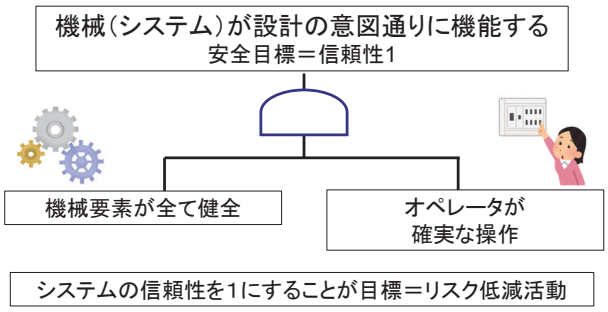

図1　機械安全の考え方（Safety-Iの仮定）

こうした考え方に基づく安全へのアプローチがSafety-Iである。すなわち、故障やヒューマンエラーなど、ものごとをうまく行かなくさせることを限りなく減らすこと（to reduce things that go wrong）[1]が、安全活動となる。

（2）標準化とSafety-I

産業現場において、Safety-Iの考え方は標準化を行きとどかせられる組み立て製造業で採用されている。

まずラインに流す製品は標準化する。その上で、生産設備の信頼性を高める。また、組み立ての正しい手順を定義し、それが実施できるような作業環境（例えば、照明条件や作業台の高さなど）をセットし、作業者には手順書を与えてそれを教育訓練し、定められた手順以外のことを行ってはいけないことを徹底する。ただし、疲労や意欲の低減により、意図せず正しい手順を果たせなくなる場合があるので、そうなる前に、休憩やリフレッシュタイムを与えるなどの使役管理を行う。

しかしながら、何らかの事情により作業者が正しい手順を実施しなければ、それはヒューマンエラーであり、製品は正しく製造されない。従ってヒューマンエラーは撲滅されなくてはならないので、なぜエラーが生じたのかの原因分析がなされ、対策を講じる再発防止活動がなされる。

（3）Safety-Iと製品安全

一般消費生活用製品の製品安全でもSafety-Iの考え方が採用される。製品構成要素の信頼性を高めると同時に、使用においては、メーカの設計者の定めた通りにユーザが「正しく」使用する限りにおいて事故は生じないと期待されることから、「正しい使用」以外の使用（誤使用）の撲滅が重要課題になる。

具体的には、

① ヒューマンインタフェースのユーザビリティの向上

② 取扱説明書や警告表示の改善、ユーザサポートの充実

③ 万一「誤使用」されても重大な事故が生じないように、フェイルセイフ設計を行う

などがなされる。

なお、一般消費生活用製品ではユーザや使用（利用）状況が多様であり、それらをメーカが容易にコントロールできるものではないことから、予想される誤使用を含む「通常の使用」を前提にした製品安全対策が、より丁寧に講じられる必要がある。

2.2　Safety-2

（1）Safety-IIの考え方

世の中、組立工場のように標準化に取り組める現場ばかりではない。というより、そうした現場はむしろ例外的であり、「生きている」産業現場の方が圧倒的に多い。自然や社会に開放された環境においての作業である。

例えば、土木建築、医療、サービスなどはそうであり、作業対象も作業条件も現場ごとに異なり、かつ、変化、変動する。身近な例では自動車運転がそうである。確かに機械としての自動車は正しく操作しなくてはならないが（シフトレバーをPに入れてからエンジンは起動しなくてはならない）、いざ道路に出たのなら、天候や路面環境、交通流などの様々な変動要素に適切に対応した運転を行わなくてはならない。こうした状況に適応した（Adjustした）運転により、安全裏に目的地に到着することができる。このとき、同じ状況であっても、運転者の対応（Adjustment）能力（レジリエンス能力という）が高ければ楽勝であるが、低ければ事故に陥る。つまり、こうした現場では、安全はレジリエンスの能力に依存することになることから、レジリエンスの能力を伸ばすことにより安全を求める必要がある。こうした安全へのアプローチをSafety-Ⅱという。ここにおいては、Safety-Iのように失敗を見つけ出しそれをつぶすという考え方をするのではなく、変化する状況において、うまくいくこと（成功）を増やすことで安全を求めていくこと（to increase things that go right）[1]になる。

(2) レジリエンス能力の構成

E.Hollnagelは、レジリエンス能力の構成として、図2を示している[1]。

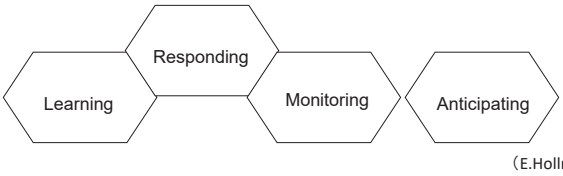

図2　レジリエンス行動を実現するための能力

このモデルを、自動車運転を例に説明する。

例えば子どもの飛び出しがあったときに、それに対応してブレーキを踏む、ハンドルを切るなどの適切な対応(respond)が必要である。それらがうまく行かないと事故が生じる。この適切な対応は事故防止への最後の砦だが、しかしそれ以前に、前方を監視し子どもの飛び出しといった状況の変化に気づくこと(monitor)、さらには、子どもが飛び出してくるかもしれないと予見し懸念していること(anticipate)が重要であり、むしろそれらが安全の鍵になっていることが多い。また、そうした予見や監視、また対応を適切に行うためには、平素からの学び(learn)が必要である。図2のモデルはこうしたことを表している。

余談だが、状況の変化に対応して安全を求めるということについてみると、入試も同じである。その場で鉛筆を転がしていては合格するはずはなく、図2のモデルに沿った行動をとることでのみ良い成果が得られる(図3)。

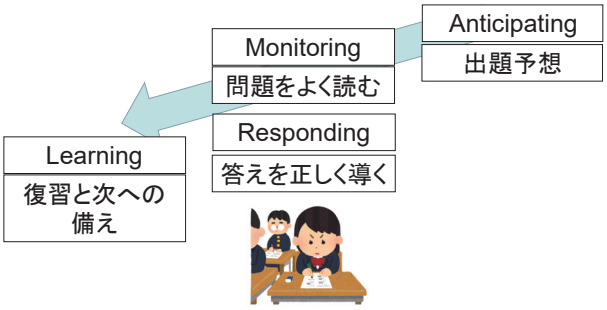

図3　入学試験もレジリエンス行動によりなされる

(3) Safety-IとSafety-IIの関係

Safety-IとSafety-IIは排他的なものではなく、また片方が他方を代替するものではない。先の自動車例に見られるように、両者が必要である。さらにいうと、道路の落石を除去することなく、腕で乗り切れというように、Safety-IでなされるべきことをないがしろにしてSafety-IIで乗り切るのはナンセンスである。要らないところでレジリエンス能力を発揮させる必要はない。つまり現実には、Safety-IとSafety-IIとは密接な関係をもっている(図4)[2]。

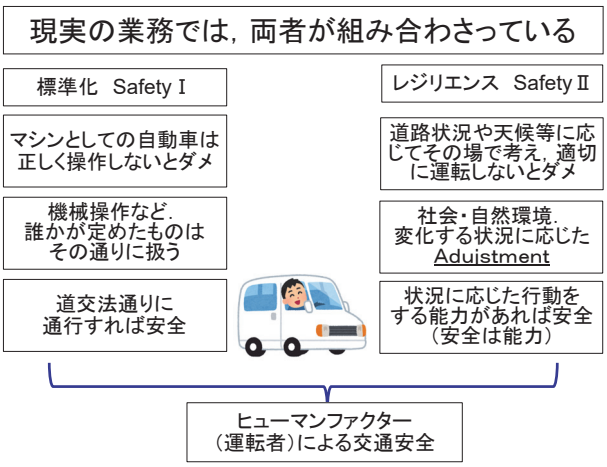

図4　Safety-IとSafety-IIは密接な関係を持つ

3. Safety-IIに対するDX技術への期待

DX技術としてIT、IoT、AIなどのいわゆるデジタル技術を想定すると、レジリエンスな行動による安全のために多くの期待が寄せられる。

① 教育訓練

Safety-Iでは、唯一のものとして存在する正しい手順を教え込むことが教育訓練の目的である。

一方、Safety-IIにおいては、あり得る現実シナリオは無数にあり、しかもあるシナリオにおいての対応の仕方も、One Best Wayが存在するというものではない。シナリオと対応の仕方の組み合わせとをすべてマニュアル化し教え込むことは不可能である。

そこで教育訓練の仕方もSafety-Iの場合とは変わり、「コンピテンシー(よりよい成果につながる行動特性)を伸ばすこと」に力点が置かれる。しかも、教育訓練コストを考えれば従来の徒弟制のように時間をかけることはできず、短期間に効率よく行う必要性がある。

シミュレーターによる訓練、行動記録映像の振り返り

など、DX技術の活用が期待される。

② レジリエンス行動への支援ツール

図2に示した各要素に対する支援を行うことで、コンピテンシーが不足している者も状況に応じた行動を行うことができる。

自動車を考えてみればよい。ルームミラー（monitor支援）、カーナビ（anticipate支援）、アンチスリップ制御装置（respond支援）などは、Safety-Iのための安全装置ではなく、全てレジリエンス行動のための支援装置である。

4. Safety-IIでの事故と製品安全

4.1 Safety-IIで起こり得る事故

Safety-Iでは正しい手順以外のことを行うことがヒューマンエラーであり、それが引き金となって事故が生じえる。正しい手順は定義されているので、エラーも、正しい手順以外の行為として事前に定義することが出来る[3]。

一方、Safety-Ⅱでは、Safety-Iと同じ図式でのヒューマンエラーは存在しない。ただし、その時点では状況に適切に対応しようとしていても、結果的に望ましくない事態が生じたときに、後知恵で後悔することはある。あえて言えば、それがヒューマンエラーということになるが、Safety-I とは異なり、あくまで後知恵である。また、何をもって望ましくない、とみなすか、ということにも依存する。初心者であれば大成功でも、ベテランにとっては大失敗、ということもある[3]。

しかしいずれにせよ、「望ましくない事態」は事故ともいえるものでもあるから、それは減らしていきたい。ではなぜ「望ましくない事態」が生じるのだろうか？

① 能力不足

その人の持つレジリエンス能力の限界を超える事態に遭遇すると、成功は保証されない。実力を伴わないまま入試に立ち向かうようなものである。

能力の限界を超える事態に遭遇すると、人は次の行動をとる。

- ● 限界能力で何とかしようとする。悪あがきである
- ● 暴走する。パニックになって支離滅裂な行動をとるようなことである
- ● フリーズする。どうしてよいか分からずに固まってしまう

② 齟齬

複数人がおり、それぞれ最善を尽くして状況に対応しているのだが、話が噛み合わないようなときに、全体としてみるとちぐはぐの行動がなされ、結果的に望ましくない事態が生じる。図5のように、それぞれが荒波を超えようと必死にオールを漕ぐのだが、そのタイミングなどが合わずに船の速力が出ず、場合によると転覆してしまうようなものであり[4]、機能共鳴型事故と呼ばれる[5]。

それぞれが局所的に良かれと思ってやっていても，
その食い違いにより，事故になる！

要するに話が噛み合わない（同床異夢）

- 前提の齟齬
- 制御の齟齬
- タイミングの齟齬
- 資源の齟齬　etc

図5　機能共鳴型事故の例え

（1）全自動機器での製品安全

全自動を謳い、判断機能が搭載されている機器では、上記のようなトラブルは起こり得る。

【全自動洗濯機での体験】[6]

- ● 洗濯ものの量に応じて水量が自動調整される。あるとき、その洗濯機の能力の限界を超える大量の洗濯物を投入したところ、最大水量で洗濯がなされ、結果的に全く汚れ落ちがしなかった（洗濯機の悪あがき）。
- ● 毛布洗いをしたところ、洗濯機がウンともスンともいわなくなった（実際にはつけ置きモード）。私はその洗濯機の対応が理解できず、電源のオンオフを繰り返す操作を行ってしまい、結果として毛布洗いが出来なかった（洗濯機との齟齬）。

（2）AIと製品安全

前述の全自動洗濯機の事例は、洗濯物の状況に応じた

作動を試みている洗濯機が起こした、ユーザにとっては望ましくない事態である。ただし、洗濯機からすれば設計者が事前にプログラムした通りの作動を行っているので、設計者の意図に反する使用をユーザが行った「誤使用」ともいえ、Safety-Iにおいての事故とも言える。

しかし、今後、AIが搭載され、状況に応じてより能動的な判断を自ら行い作動する機器が本格的に出現してきた場合、AIが何をどの程度学習したのかにより、同じ状況を与えても機器により作動の仕方が変わる可能性がある。つまり設計者の意図が無存在に機器が挙動する。そのとき、望ましくない事態が生じた場合に、それを製品安全としてどのように考えるかは、今後の課題ではないだろうか。

ペットを考えれば良く、同じ母犬からから生まれた仔犬も、飼い主次第で賢くもなり、狂暴にもなる。その凶暴な犬が第三者に噛みついた場合には、飼い主責任が問われる(噛みついた犬にも母犬にも責任はない)。

同様に、育て方次第でAIが望ましくない事態をもたらすことはあり得る。その責任はAIにはとりようはなく、またメーカ責任でもない。そうなるとAIを育てたユーザ責任となるのだろうか。

これらを含め、AIのあるべき姿はAI倫理[7]において議論されていることであるが、今後の製品安全の責任分担という点で議論が必要になると思う。

ロボットの安全も、プログラム通りに作動する産業用ロボットの場合と、自然、社会環境下で能動的に挙動する生物型ロボットでは異なる部分があるように思う(図6)。

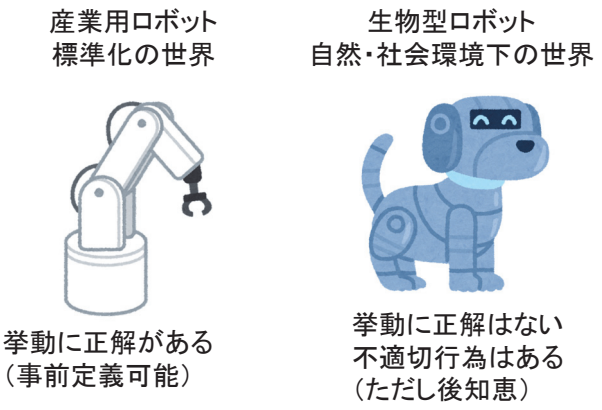

図6　産業用ロボットと生物型ロボットの違い

5. おわりに

消費生活用製品においては、今までの製品は受動的という言い方が出来る。つまり、ユーザに利用されるがままであるから、お願いだから「正しく使ってね」というSafety-I的アプローチが基本であったと言える。そのため、従来の製品安全では、その製品はどのように使えばよいのかを使用者に説明する責任が求められ、それに呼応して使用者は正しい使用を行うという図式のもとに成立している。この図式のもと、「どのように使えばよいのか」について、通常の注意や努力を超える使用をユーザに求める製品は人間工学上の設計の欠陥を有し、またその説明が不足している場合は、説明の欠陥を有することになるから、利用状況を丁寧に把握した上での設計対応や使用説明が求められてきた。

一方、これからは、より能動的な製品の出現が考えられる。つまり、自律的に状況に応じた挙動をする機械、しかも使い込まれるほどに学習する機械が出現する。この場合の製品安全はSafety-II的アプローチが求められると考えられる。能力を超えた事態でどう挙動させるか、機能共鳴型事故をいかに防ぐかなどであり、さらには、如何に賢く育てるか、ということも、望ましくない事態が生じた場合の責任所在との関係で課題になる。

参考文献

[1] E.Hollnagel(北村正晴・小松原明哲監訳)、Safety-I & Safety-II 安全マネジメントの過去と未来、海文堂、2015.

[2] Komatsubara, A.Development of Resilience Engineering on Worksites, (in Advancing Resiliet Performance, ed by Christopher P.Nemth and Erik Hollnalgel), Springer 2021.

[3] 小松原明哲、ヒューマンエラーの考え方―人間工学の立場から(甲斐克典編、医療安全と医事法[医事法講座第11巻　第1章])、信山社、2021.

[4] 小松原明哲、安全人間工学の理論と技術　ヒューマンエラー防止と現場力の向上、丸善出版、2016.

[5] E.Hollnagel(小松原明哲監訳)、ヒューマンファクターと事故防止・”当たり前”の重なりが事故を起こす、

海文堂、2006.
[6] 小松原明哲, 人にやさしいモノづくりの技術　人間生活工学の考え方と方法, 丸善出版、2022.
[7] 高橋利枝、国連『AIのある未来』:人を幸せにする持続可能な社会の創造に向けて、人間生活工学24(1)、pp1-7(一社)人間生活工学研究センター、2023.

略歴

小松原 明哲 (こまつばら あきのり) 博士(工学)
日本人間工学会認定人間工学専門家
専門:人間生活工学、安全人間工学
1980年　早稲田大学理工学部工業経営学科卒。産業医科大学医学訪問研究員、金沢工業大学教授を経て、
2004年　早稲田大学理工学術院創造理工学部　経営システム工学科教授