alliance to counter crime online

Summary: We encourage Ofcom to shift its approach to regulating online platforms from the current "risk management" approach to a "product assured safety management" approach. In this white paper, we explain why this change would represent more than just a subtle semantic shift. In fact, it would position Ofcom to encourage safety, rather than respond to risk, and to stop problems before they emerge, rather than cleaning them up afterwards. We assert that this approach would significantly improve Ofcom's regulatory effectiveness and certainty, reduce costs and uncertainty for industry, and also result in improved safety and performance in the online realm for all participants – in other words, a regulatory hat-trick.

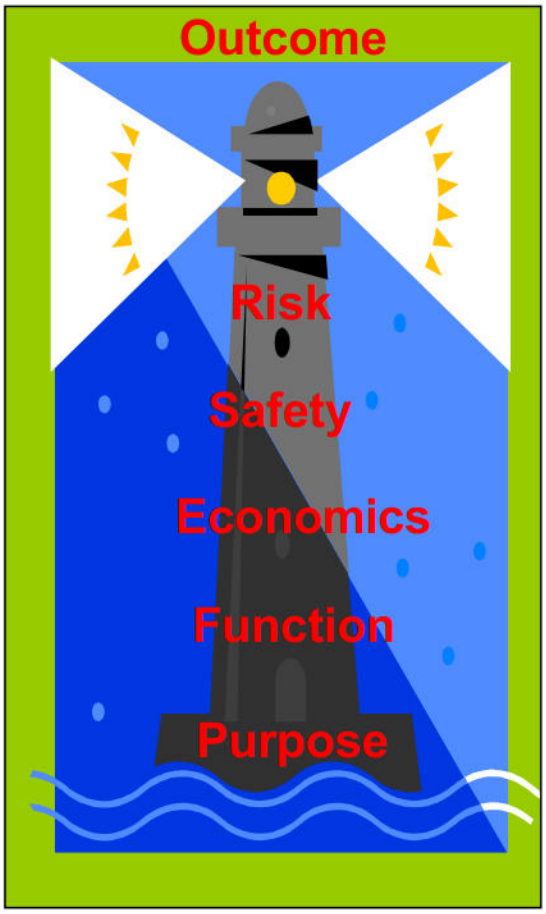
Background: Almost a half-century ago, the Health and Safety at Work etc Act 1974 (HSWA74) came into force after the Aberfan disaster of October 1966, in which 116 children and 28 adults were buried alive in a collapsing coal slag heap. Then as now, with the Royal Assent of the Online Safety Act, the British government adopted legislation that was intended to retroactively regulate a set of products already in wide commercial use, and that it had come to realize had potentially deadly safety risks for the public. Regulators tasked with enforcing HSWA74 equipped Health & Safety Inspectors with codes of practice for machine safety, which set standards for how good safety devices and machines operated. Knowing and responding to harms was not enough, inspectors needed to understand how to direct manufacturers to develop machinery that did not harm people in the first place. This implicitly required a practical understanding of what safe design would look like. Importantly, safe design does not mean an absence of danger. Rather, it means that each product must be evaluated as "fit for purpose" when properly in use. For example: a factory tool for cutting aluminum sheets will inevitably pose certain dangers, and therefore must be designed with certain requirements and user regulations to best ensure the safety of the user. As another example, automobiles will inevitably be involved in accidents, but regulations require things like seat belts, air bags, antilock brakes, and collision-warnings systems to both prevent harm arising and to reduce the consequences of those accidents.

> **HSWA74** has as its objective: "*securing the health, safety and welfare of persons at work, [and] protecting others against risks to health or safety in connection with the activities of persons at work ….* "

The Challenge: When HSWA74 was adopted, both workplaces and products tended to be compact physically and to operate as independent units, with little joining them together. In 2023, the challenge is to regulate activity in cyberspace, a borderless and massively interconnected space that is far more complex to assess and address than a traditional brick-and-mortar factory, in particular since the UK is but a small part of that global online space. Until now, tech companies have been permitted to release products into the global market in a near absence of regulatory standards, and also have been allowed to more or less define for themselves what "good enough" looks like in addressing harm on their platforms. Now it is incumbent on Ofcom to figure out how to regulate both large and small tech companies whose platforms are available for use in the UK. Although enforcement challenges exist, the same safety management approach to regulating a factory or physical device can – and should – be applied to regulating Internet platforms. We assert that a product assured safety management regime will do more to reduce risks to users than a regime that simply produces regular reports on risks.

> **OSA2023** has as its objective: "*making the UK the safest place in the world to be online.* "

The Solution: We propose a purpose-to-outcome "Lighthouse" model for evaluating online products. Under this model regulators would implement a multi-step process to identify the Purpose of each product, and determine whether or not it achieved that Purpose. Regulators would determine that each product was "fit for purpose" (FFP) and "safe" (S), with risk management becoming a part of the process, not the process itself. It's important to recognize that "Safe" is a state that comes from a product or service being BOTH fit for purpose AND not causing damage to people, property, or environment when properly in use. A product or service can be FFP and not entirely Safe. A product or service cannot be Safe, if is not FFP.



1. Outcome: The intended outcome of a lighthouse is to provide a guiding light and reliable reference navigation point. Likewise, the ultimate goal of the Lighthouse model of regulation is determining whether the as-delivered product is navigated towards successful achievement of its intended Purpose.

2. Purpose: The foundation of the Lighthouse model is determining the Purpose each product aims to achieve. A social media platform aims to connect people and businesses within a friendly forum where they can share messages, photos, audio and videos. An eCommerce platform aims to facilitate its users to buy and sell items or services. A gaming platform aims to provide an online space for people to play, and in some cases to also communicate and transact. These are but a few examples; there exist a great number of Internet platforms, and the purposes for which they exist grow more varied with each year and technological advance. Design teams thus become vital to establishing and clarifying purposes that will satisfy regulatory requirements. Ofcom could require tech companies to provide Purpose Statements for each of their products and service to enable the process of determining if they are FFP. Any companies that fail to provide Purpose Statements could be deemed to be in breach, and could be required to undergo an enforcement intervention that would also benefit the company.

3. Function: This multi-step aspect of the process maps out and clearly details the mechanics for achieving the product's Purpose, as well as the various functionalities of the product. The function must align with the Purpose, or else the Outcome will be failure. Engineers will be critical to implementing a meaningful interpretation of the Design requirements that satisfy the Purpose. Regulators must have the opportunity to "look under the hood and kick the tyres," in particular when working with complex artificial intelligence, end-to-end encryption, algorithmic recommendation, and human language processing tools, in order to understand precisely how they operate, and their limitations.

4. Economics: In this step, regulators ensure that ongoing Time, Money & Effort will be sufficient to achieve the intended Purpose using the proposed function(s). At this stage a product needs to have all three elements (purpose, function and economics) in a symbiotic relationship with each other such that the running cost is funded by the product in operation. A product is not viable if it continues to require external funding and input just to keep it going. Completing this 3-parameter viability check provides advance indications of potential future service failure.

5. Safety: Once regulators have established the viability of a platform's integrated purpose, function, and economics, they can work to assess the core stability of that product, and create a safe maneuvering limitation diagram (MLD) for that product in service. The product will already have had safety features integrated into it during prior phases; this phase will ensure product viability by evaluating and assuring

that safety requirements cascade and/or radiate throughout every facet of the design, manufacture, and operation. Ofcom could regulate tech companies to operate under a Code of Conduct for designing new products, setting standards and expectations for products to meet their purpose (a bit like a car brake performance check), protect users from harmful and illegal content, and identify steps tech companies could implement to prevent illicit actors and predators from abusing their product(s). A key aspect of this process would require software engineers who design these products to be educated in compliance, and to receive routine training to maintain up-to-date knowledge and reports about the ways that illicit and malign actors and predators use and abuse online services. Engineers would be required to return to the question of "does product meet purpose" after any new build or remedial work occurs.

6. Risk Assessment: Virtually any product – be it in the physical world or in cyberspace – bears some risk. The risk assessment aspect of the regulatory process is aimed at ensuring that uncontrollable external factors do not de-stabilise the core functions and purpose of the product. As each product will be slightly distinct, companies must create a unique risk evaluation matrix (that expresses the company's tolerance for risk) for each new product released onto the market, just as is the case for manufacturers of automobiles, medical devices, and other technologies. In the digital age, it is especially important that manufacturers of online products consider the various ways their products may be prone to being repurposed for malicious intent. While creating unique risk assessments for every online product already on the market will no doubt create a significant workload for Ofcom at the outset, it will also assist regulators to identify the difference between low-, medium-, and high-risk products, and will pave the way for more streamlined evaluation processes in the future. Importantly, just as Health & Safety regulators who enforce HSWA74 must evaluate risks inside and outside factories, any risk evaluation of online platforms must examine unintended consequences, both good and bad, occurring both on and off the platforms being assessed.

When Your Safety Becomes My Danger: One of the most challenging aspects for Ofcom will be regulating products and services, chiefly E2EE and disappearing content, that may be favored by some constituencies for safety reasons, but which also create dangers for other groups. There is no one-size-fits-all solution to this dilemma, although we assess that regulations must always be geared towards protecting the most vulnerable constituencies.

Proportionate Measures: OSA2023 currently provides an incoherent definition of how it will enforce 'proportionate measures' (Vol 3 and Annex 10) depending on the size of a platform. We assess that measures should be proportionate to the risk taken and harm caused, not to the size of the platform. The UK's Health & Safety Executive has a long history of conducting Impact Assessments and Cost Benefit Analyses to ensure that the cost of mitigation is not disproportionate to the benefit of risks that do not require to be absolutely controlled. This model could also be repurposed for the online realm.

Typologies: There are important parallels between regulating tech and the way in which the banking industry is regulated. As with Internet platforms, banking services are mainly provided by large companies serving millions of users around the globe. Those users have reasonable rights to privacy, although it's widely accepted that bank services have been abused for illicit purposes, such as money laundering and fraud. Banks are therefore legally required to monitor their systems for illicit conduct and report suspicious activity to law enforcement. To comply with user privacy regulations, banks and law enforcement have developed "typology reports" as a way to share information. Typology reports describe methods and patterns of criminal actors, and can be updated regularly to keep compliance teams up to date. Banks utilize these reports to search for illicit activity within their systems, and reduce their exposure to criminal activity. ACCO has proposed that similar reports could be produced, perhaps in collaboration with law enforcement and civil society groups, to respond to and reduce illicit activity on the surface web.

Enforcement & Ratings: Ofcom has a huge regulatory task ahead, and enforcing UK laws on platforms operated from outside its territory will be a challenge. Britain may elect to ban certain risky services, but technologies like VPNs and satellite Internet services will provide easy work-arounds for sophisticated users. In order to protect vulnerable constituencies, in particular children, it may be useful to create a ratings system, similar to those applied to film and music, so that parents can easily evaluate relative platform risks. It may also be advisable to require the production of online safety devices that would be inserted at system level to inhibit work-arounds.

Conclusion: By creating a common thread that stitches together the design, manufacture, testing, and evaluation of online products, companies and regulators alike will be able to reduce harmful outcomes, and more quickly identify what "good" looks like. Implementing a purpose-to-outcome process will also illuminate circumstances when the thread is broken or missing, making it easier for regulators and the private sector alike to target action before any harm can escalate out of control. We urge you to consider this approach, which we believe will significantly improve Ofcom's regulatory effectiveness, increase the certainty that products are utilized as intended, and even reduce costs for industry, while improving safety conditions in cyberspace for all British Internet users. We'd be delighted to help you design and execute a plan to implement it.

Authors: *Gretchen Peters is Executive Director of the Alliance to Counter Crime Online. Peter Hanley is a former UK regulator with Health & Safety Executive (HSE) with years of experience regulating a wide range of manufacturing industries from food, dairy farming, factories and transport to nuclear support services.*