
This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

http://eprints.gla.ac.uk/130826/

Deposited on: 01 November 2016
Role of Regulators in Safeguarding the Interface between Autonomous Systems and the General Public

Chris. W. Johnson DPhil, School of Computing Science, University of Glasgow, Glasgow, UK.

Keywords: Autonomous systems, Drones, RPAS, Safety, Cyber Security.

Abstract

Regulators play a critical role in the commercial exploitation of new technologies. They protect the public when market competition might persuade companies to take undue risks. At the same time, it is essential that regulatory authorities do not kill innovation by imposing inappropriate rules or by retaining previous requirements that make little sense in the light of technical innovations. These tensions are apparent in the introduction of autonomous and semi-autonomous systems, across a range of industries. ‘Regulatory lag’ has starved companies of the strategic guidance that is necessary to make informed decisions about acceptable levels of safety and security for the integration of these technologies. This paper argues that existing product-based, process-based and performance-based approaches to regulation threaten the safe and secure exploitation of new markets. In contrast, we advocate a Competent, Anticipatory, Self-Reflective approach, which places performance requirements on the regulator rather than on the markets they protect.

Introduction

The United States National Highway Traffic Safety Administration (NHTSA) (Ref.3) identifies a spectrum of autonomy:

- **Level 0**: The driver completely controls the vehicle at all times;
- **Level 1**: Individual vehicle controls are automated, such as electronic stability control or automatic braking;
- **Level 2**: At least two controls can be automated in unison, such as adaptive cruise control in combination with lane keeping;
- **Level 3**: The driver can fully cede control of all safety-critical functions in certain conditions. The car senses when conditions require the driver to retake control and provides a "sufficiently comfortable transition time" for the driver to do so. The Tesla Model S provides an example of this level of automation;
- **Level 4**: The vehicle performs all safety-critical functions for the entire trip, with the driver not expected to control the vehicle at any time. As this vehicle would control all functions from start to stop, including all parking functions, it could include unoccupied cars. Google’s Self-Driving Car Project illustrates full autonomy (Ref.1). These lack the physical controls that might otherwise support direct human control.

In aviation, the FAA has introduced similar distinctions in their Notice of Proposed Rulemaking and subsequent requirements for the Operation and Certification of Small Unmanned Aircraft Systems (RIN 2112–AJ60). A Remotely Piloted Airborne System (RPAS) may operate under the direct control of a ground-based pilot equivalent to NHTSA levels 1 and 2. It may also operate under automated control for specific operations providing that the pilot retains ‘line of sight’ with the vehicle; equivalent to levels 3 and 4. However, full autonomy is not permitted without specific waivers, for example ensuring that operation is restricted to carefully controlled, experimental contexts well away from controlled airspace. Other industries are also exploring acceptable levels of autonomy in safety-related applications from the autonomous monitoring of ‘Smart’ power distribution grids (Ref.4) through to robotic surgical systems (Ref.5).

New generations of autonomous systems create significant concerns for public safety. In particular, it is unclear how to integrate these applications with more conventional systems under direct human control (Ref.6). For example, the testing of automated vehicle technologies is only possible in the UK if a test driver is present and takes responsibility for the safe operation of the vehicle and for the safety of other road users. There are further concerns:
• *Maintaining situation awareness.* It is hard for human operators to monitor semiautonomous systems, when little direct intervention is required. Boredom and inattention can lead to significant safety concerns. Even in fully autonomous systems, it is important for operators to respond to degraded modes of operation – for instance alerting other airspace users to potential problems with RPAS (Ref.7);

• *Supporting manual intervention.* Over time, manual control skills can be eroded so that even when operators maintain adequate situation awareness, they lack the necessary skills to assume control of a semiautonomous system. Bainbridge provides an eloquent review of the ‘Ironies of Automation’ – the more successful automation becomes then the less prepared operators that will be to respond to failure modes (Ref.8);

• *Maintaining cyber security.* It is important that operators and engineers identify potential vulnerabilities and mitigate threats to the cyber security of autonomous and semiautonomous systems (Ref.9). The lack of direct manual control increases the possibility that attackers will disrupt or redirect the operation of these platforms and their associated communications links;

• *Lack of appropriate historic data.* In conventional applications, companies can draw on a wealth of experience to guide key decisions in the design and operation of complex systems. Unfortunately, the novelty of autonomous systems means that we have relatively little expertise in appropriate means of integration with conventional systems (Ref.10). In particular, it is very unclear how best to organize the interactions that might arise between multiple autonomous and multiple conventional systems.

This is not an exhaustive list of the concerns that arise from the introduction of autonomous systems across many different industries. In the face of such uncertainty, commercial organizations must identify acceptable levels of safety and security when seeking to exploit new market opportunities. This raises further questions; can the public place an appropriate level of trust in the safeguards provided by companies competing to develop new markets without some form of independent guidance? There is also a concern that commercial pressures will influence the subjective interpretation of ‘acceptable’ levels of risk.

### Why Regulate Autonomous Systems?

Governments can exploit a number of mechanisms to increase public confidence and at the same time promote the development of autonomous and semiautonomous systems (Ref.11).

**Limits of Market Forces:** Market forces reduce the need for regulation in the design and operation of autonomous systems. If companies fail to ensure the safety and security of the general public then customers will not continue to buy their products. This ignores the limitations of ‘imperfect information’. In other words, customers may be unaware of the high accident and incident rate associated with existing systems. Even when they are aware of the risks, safety information can be obscured by coverage of individual success stories. These concerns have been identified in studies of military RPAS; many nations have sustained significant financial and operational losses as they rush to exploit new capabilities (Ref.12).

Third party effects also limit the use of market forces to ensure safety and security. These arise when the victims of an accident are not involved in the demand or supply of autonomous systems, for example, bystanders at a sporting event that is being filmed using an RPAS. In such circumstances, market forces will not limit the supply of potentially dangerous products because the operators do not perceive the risks they create for the general public. It is for this reason that the FAA requires the registration of lightweight drones. This enables legal redress against the individuals and companies that endanger safety.

**Limits on Insurance and Tort:** Tort law helps to mitigate third party effects; injured parties can sue the operators and manufacturers of autonomous systems. Financial costs and punitive damages can be awarded after an incident. Unfortunately, Tort is retrospective. Litigation takes place after an accident has occurred. Some US states now require a surety bond of $5 million for the operation of autonomous vehicles (Ref.16). However, the manufacturers and operators of autonomous systems can be persuaded to offset the costs of litigation and surety by purchasing insurance (Ref.13). Insurance can have a proactive effect when actuaries reduce premiums for operators and manufacturers that demonstrate strong security/safety cultures.
Litigation and insurance do not provide a panacea. In particular, it is often hard for underwriters to calculate appropriate premiums when they cannot easily assess their future exposure based on past risks. The consequent cost of any policy may kill the development of a fledgling industry, especially if all potential users are compelled to carry such insurance.

Concerns for the Regulation of Autonomous Systems

Given the concerns over the unhindered use of market forces and of litigation to support the safe development of autonomous and semiautonomous systems, governments have accepted the need for regulation (Ref.14). However, there is huge uncertainty over appropriate forms of intervention with inconsistent approaches being adopted both across and within industries around the globe.

Inconsistency: It is hard to exaggerate the diversity of approaches to the regulation of autonomous systems. The rules are different from one country to the next. This is illustrated by Table 1, which contrasts the different requirements for RPAS operation across North America. These differences arose even when the two countries tried to reflect the requirements of their neighbor. The classes of airspace mentioned here are derived from the International Civil Aviation Organization – for example, air traffic management does not control class G operations and any clearances are advisory. In contrast, class E air space is controlled for aircraft operating under instrumented flight (IFR) but uncontrolled if using visual flight rules (VFR). Further details are provided in the ICAO Annex 11.

<table>
<thead>
<tr>
<th>PROVISION</th>
<th>CANADA</th>
<th>USA SMALL UAS</th>
<th>USA MICRO UAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>&lt; 4.4 lbs (2 kg)</td>
<td>&lt; 55 lbs (25 kg)</td>
<td>&lt; 4.4 lbs (2 kg)</td>
</tr>
<tr>
<td>Maximum Altitude</td>
<td>300 ft</td>
<td>500 ft</td>
<td>400 ft</td>
</tr>
<tr>
<td>Airspace Limits</td>
<td>Class G only</td>
<td>Class G and E. Class B, C, D with ATC permission</td>
<td>Class G only</td>
</tr>
<tr>
<td>Distance From Obstacles</td>
<td>100 feet laterally from structures, 100 feet from people</td>
<td>Operation prohibited over any person not involved in operation</td>
<td>Flying over any person is permitted</td>
</tr>
<tr>
<td>Operational Area Extension</td>
<td>No</td>
<td>Yes, from a boat</td>
<td>No</td>
</tr>
<tr>
<td>Autonomous Operations</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Aeronautical Knowledge Required</td>
<td>Ground school</td>
<td>Knowledge test</td>
<td>Self-certification</td>
</tr>
<tr>
<td>FPV Permitted</td>
<td>No</td>
<td>Yes, if you can also visually see UAS</td>
<td>No</td>
</tr>
<tr>
<td>Operator Training</td>
<td>Ground school</td>
<td>Not required</td>
<td>Not required</td>
</tr>
<tr>
<td>Visual Observer Training</td>
<td>Ground school</td>
<td>Not required</td>
<td>Not required</td>
</tr>
<tr>
<td>Operator Certificate</td>
<td>Not required</td>
<td>Required, must pass basic UAS aeronautical test</td>
<td>Required, no knowledge test</td>
</tr>
<tr>
<td>Preflight Safety Check</td>
<td>Required</td>
<td>Required</td>
<td>Required</td>
</tr>
<tr>
<td>NearAirport Operations</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Congested Area Operations</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Liability Insurance</td>
<td>Required, $100,000</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Daylight Only</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Aircraft Made of Frangible Materials</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1: Comparison of US and Canadian RPAS Requirements

Regulatory guidance on autonomous vehicles in Europe and in North America also varies between individual regions within the national legal systems (Ref.15). In the USA, individual state laws vary significantly and “no state has fully determined how existing traffic laws should apply to automated vehicles” (Ref.16). Four states have explicit provisions supporting the introduction of this technology. Fifteen have rejected bills related to automated driving even though the US Department of Transport and the National Highway Traffic Safety Administration (NHTSA) remain committed to supporting the introduction of these technologies.

1 http://spectrum.ieee.org/automaton/robotics/drones/faaproposedcommercialdronerules
Similar inconsistency can also be seen within Europe. For example, the Germany Federal Highway Research Institute has argued that fully automated vehicles do not comply with existing traffic law. Each Federal state can grant exemptions from the German Road Traffic Licensing Regulations to allow tests ‘provided there is a driver in the driver’s seat who has full legal responsibility for the safe operation of the vehicle’ (Ref.16). In France, specific zones have been established for testing, including changes to driver training. There is also provision to allow ‘large-scale’ testing of self-driving cars and trucks. Sweden has followed a similar approach, allowing tests as part of the Volvo ‘Drive Me’ project in restricted areas around Gothenburg.

These differences arise partly from the novelty of these systems and the lack of international agreements about acceptable parameters for the safety and security of these applications. They also stem from international competition as countries seek to ensure that their industries are not left behind in the race to develop new technologies. However, regulatory inconsistency has a profound impact especially on international companies who must conform to the local requirements that create barriers to the development of a global market place.

**Regulatory Competence:** The fiscal crisis has prevented many government agencies from hiring and retaining staff with expertise in emerging technologies. Limits on promotion and salary caps can persuade the best engineers to leave government service and join ventures developing autonomous systems. In such circumstances, it is hard for regulators to go beyond generic guidance and provide sustained, technical support for the companies developing autonomous applications. This is entirely appropriate when commercial organizations are concerned to protect intellectual property rights in a dynamic market place. However, the limits of regulatory competence create significant concerns for the future. This can be illustrated by existing confusion over the introduction of Artificial Intelligence and Machine Learning within safety-critical, autonomous systems. Existing standards, including IEC61508 and ISO26262, exclude the use of such technologies at higher levels of integrity because it is difficult to prove that they will not ‘learn’ potentially dangerous behaviors.

The Vienna Convention on Road Traffic requires that ‘every moving vehicle or combination of vehicles shall have a driver’ and that ‘every driver shall at all times, be able to control his vehicle’ (Ref.16). Some have taken this to be a barrier to the introduction of automated vehicles. However, recent announcements by the NHTSA legal counsel imply that the machine learning systems embedded within autonomous vehicles could be considered as a driver under federal law (Ref.17), leading the way to the use of AI techniques as a primary means of control on the public highway. We argue that such contradictions between the legal and technical arguments, further demonstrate the need for fresh thinking in the field of safety regulation.

**Regulatory Lag:** In 2012, the United States Congress mandated the FAA to issue specific guidance on the regulations governing the introduction of RPAS into the national airspace. The intention was to provide a comprehensive framework that would replace the ad hoc exemptions that supported the commercial operation of drones on a casebycase basis. However, the anticipated deadlines were missed so that by 2015 more than 1,000 companies had been issued FAA333 exemptions. This delay can be justified – for example by arguing that the use of waivers enables the FAA to gain sufficient operational expertise to inform the subsequent development of regulatory requirements. In contrast, the UK government has advocated a Code of Practice to promote safety and set clear guidance in the use of autonomous vehicles. It is argued that a Code of Practice ‘will be quicker to establish, more flexible and less onerous for those wishing to engage in testing than the regulatory approach being followed in other countries, notably in the US. Failure to follow guidance in a Code of Practice would be a clear indicator of negligence… Those wishing to conduct tests are not limited to the test track or certain geographical areas, and do not need to obtain certificates or permits’ (Ref.16). The Code of Practice would be subject to frequent revision without the delays associated with legislative changes or with the formal consultations associated with existing regulatory practice. However, this flexibility carries considerable uncertainty when companies must ensure compliance with successive changes in the underlying provisions in these documents.

These examples drawn from autonomous and semiautonomous systems are symptomatic of wider concerns over ‘regulatory lag’. In particular, many industries continue to complain about the lack of detailed cyber-security guidance even though regulatory agencies clearly acknowledge the growing threat to those companies and to the public (Ref.18). Regulatory lag creates significant problems for government, for industry and for the general public:
• Elevated Risk. Without specific guidance, there is a danger that companies, which have not adopted appropriate mitigation techniques when deploying autonomous systems, may place the public at increased risk.

• Barriers to Innovation. Without specific regulatory guidance, companies can be dissuaded from entering a market with the danger that their products and services may fail to meet subsequent requirements.

• Technological Flight. Without specific regulatory guidance, there is a danger that companies will choose to move to other jurisdictions where there is greater certainty in the requirements that they must meet before being allowed to operate.

• Legal uncertainty. Without specific guidance, companies cannot easily assess whether or not they would be liable from any litigation following an adverse event.

Many of the causes of regulatory lag have been introduced in previous sections. They include the difficulty in retaining technical competence, concerns over premature regulation, political pressure to stay out of emerging markets etc. However, there are more systematic causes that arguably stem from outdated regulatory practices dating back to an era when there was a far slower pace of technical innovation and when there were few truly global markets (Ref.11).

New Approaches to the Regulation of Autonomous Systems

The delays in providing regulatory guidance for integrating autonomous and conventional systems and the contradictions between technical standards, such IEC61508, and existing commercial practice are symptomatic of deeper problems in the regulation of complex, sociotechnical systems. In particular, there is an urgent need to move beyond product, process and performance based regulation to devise ways of protecting the safety and security of the public at a pace that at least approximates to the speed of recent technological innovation.

Product vs Process vs Performance Based Regulation: In order to set any new proposals in context, it is important first to sketch the three predominant approaches to safety regulation. Product based techniques specify standards against which one can measure particular artifacts. For example, EN ISO 20346:2004 provides means of determining whether or not a shoe provides sufficient protection for use in safety-related industries. A specific item of footwear can be inspected to determine whether it includes an appropriate steel toecap.

Unfortunately, this product-based approach works less well for more complex systems. In particular, it cannot easily be applied to the logical abstractions within software. Standards such as DO178C and ISO26262, therefore, focus less on artifacts and more on the processes that are used – for example to identify requirements or to allocate validation and verification resources. This maximizes regulatory resources because auditors do not need to consider millions of lines of source code. Instead, they can focus on assessing the quality of the higher level processes that contributed to the development of that code.

The limitations of process-based regulation include the need for regulators to understand the techniques that are used within complex application domains. It can be hard for regulators who are not practitioners to remain up to date with leading software development practices. There are other concerns, for instance when regulators are too closely associated with the companies that they regulate. It is for this reason that many countries now employ multimodal regulators. These are not specialists in any one domain. Instead, it is argued that regulators should develop expertise in regulation that can be applied across multiple industries.

The deployment of multimodal regulators often supports performance-based regulation. Rather than focusing on the properties of a particular product, or on the processes that were used to develop that product, the focus is on monitoring Key Performance Indicators (KPIs) for products and processes. It is important to avoid retrospective KPI’s, where we can only identify concerns after an accident has occurred. Performance based regulation focuses on leading indicators that identify potential risks before an accident occurs. As we have seen, however, it is far from clear how to apply performance based regulation to innovative and dynamic industries where we have little track record of integration – for instance, between autonomous and conventional systems. It is hard to identify
appropriate KPIs and the erosion of detailed technical competence associated with multimodal regulation has arguably compounded concerns over regulatory lag.

**Competent, Anticipatory, Self-Reflective (CAS) Regulation:** A range of techniques has been used to combat the problems of regulatory lag. The FAA has issued waivers that permit the violation of existing legislation over the years that have been required to iteratively refine regulatory guidance in consultation with dozens of stakeholder organizations. The ad hoc nature of this approach has triggered sustained criticism from industry and from politicians. In contrast, the UK government have advocated the use of less formal Codes of Practice. These can be incrementally refined but lack legal force; violations may or may not be viewed as negligent during subsequent litigation. Companies also face significant risks when complying with Codes that are frequently revised.

In contrast, we would advocate a fresh approach to the regulation of complex and dynamic industries based on three guiding principles:

1. **Competent Regulation.** The integration of autonomous and conventional systems has revealed the pressing need for technical competency. This might seem self-evident. However, many countries now use multi-modal agencies where staff are not drawn from the industry they regulate. This increases independence because regulators do not work with previous employers. Their competence is in the application of regulatory requirements rather than in their detailed knowledge of domain-dependent systems engineering. Multi-modal regulation creates particular concerns for new markets; where safety relies on technical innovation. For example, the NHTSA’s argument that autonomous systems can fulfill the safety requirements normally associated with a human driver has enormous technical implications for the application of existing standards such as ISO26262 and IEC61508. We urgently need regulators who understand these implications and can provide companies with guidance on how to ensure that AI and machine learning techniques are acceptably safe;

2. **Anticipatory Regulation.** Again it might seem self-evident that regulators should try to stay ‘ahead of the curve’. However, the delays in integrating autonomous and conventional systems shows that this is not always achieved. Regulatory lag arguably occurs because civil servants habitually seek comfort in lengthy consultation processes. This is justified given the strong political disincentives for market intervention. However, regulators often fail to catch up when safety concerns arise. Many agencies struggle to enforce acceptable modes of operation after the public has purchased thousands of mass-market drones. In contrast, we would argue that future regulation should be anticipatory based on proactive market leadership. It is often argued that safety and security have to be designed-in from early in the development cycle. It is, therefore, little surprise that regulatory support is often ineffective when drafted months and years after the deployment of complex, autonomous systems;

3. **Self-Reflective Regulation.** Our emphasis on technical competence and anticipatory regulation will only be successful if regulators learn from their own successes and failures. Many regulatory agencies are deliberately excluded from commissioning this research; from questioning their own practices. In contrast, we would stress the need for regulatory self-reflection. There is little systematic evidence to show that existing standards, such as ED-153 or DO-178C, lead to safety improvements. Self-reflective regulation extends the concepts of anticipation and participation to develop an evidence-based approach to the validation of regularity intervention. Brevity prevents a detailed analysis of the methodologies that might support such studies, however, NASA provide a template; comparing the work of multiple contractors using DO-178A to implement the same functional requirements. The aim of this early work was to determine the level of support provided through the use of DO-178 in the software development process (Ref. 11). The UK Health and Safety Executive provide a further example when they question the causes of accidents involving systems that were developed under IEC 61508 (Ref. 12). At the heart of this work is the idea that the analysis of accidents, which will always be possible even under the best regulatory regime, can also be used to improve standards and other forms of government intervention. Both of these research projects were completed in the early 1990s but perhaps it is time to resurrect these studies and understand the reasons why new industries are faced with regulatory lag in the exploitation of future markets.
Conclusions

We have argued that ‘regulatory lag’ starves companies of the strategic guidance that is necessary to make informed decisions about acceptable levels of safety and security in autonomous and semiautonomous systems. Existing approaches to product-based regulation cannot easily be applied to software intensive systems. Process based techniques often require careful involvement in commercial development practices that cannot be sustained when regulators lack technical competence in particular domains. Performance based regulation fails when it may take years before we can identify appropriate leading indicators that anticipate potential accidents. In the meantime, the ad hoc use of waivers and the drafting of temporary codes of practice provide companies with few guarantees that their investments will meet subsequent regulations.

In contrast, we have stressed the need for Competent, Anticipatory, Self-Reflective (CAS) approaches. Technical competency is required because we do not believe multimodal regulation is capable of providing the detailed guidance that is required, for example to ensure that AI and machine learning algorithms provide acceptable alternatives to human intervention. Anticipatory regulation is required because it is increasingly difficult to ensure that companies and the general public meet more stringent operating requirements after they have become use to semi or unregulated environments. Finally, Self-reflective regulation extends the concepts of anticipation and participation to develop an evidence-based approach to the validation of regularity intervention. Many regulatory agencies are deliberately excluded from commissioning this research; from questioning their own practices.

Further work is required to analyze the causes and mitigations for regulatory lag across a range of different industries beyond the specific focus of this paper on autonomous systems. Some initial studies have been published (Ref.14, 15). However, work in this area remains controversial given the sensitivity of the topic and difficulty of identifying appropriate methods for validating alternative approaches to regulatory intervention. It is clear, however, that while many previous papers in ISSC focus on specific tools and techniques for improving safety and security, only a very small number consider the regulatory frameworks that help to determine the application of those approaches within specific industries.

References


17. NHTSA, Letter to Google from the NHTSA Chief Counsel (Paul Hemmersbaugh), Washington DC, USA, February 2016.


Biography

Chris Johnson, DPhil, School of Computing Science, University of Glasgow, Glasgow, Scotland, G12 8RZ, Scotland, U.K., telephone – +44 (141) 3306053, facsimile – +44 (141) 3304913, email – Johnson@dcs.gla.ac.uk.

Chris Johnson is Professor and Head of Computing Science at the University of Glasgow in Scotland. He leads a research group devoted to improving the cyber-security of safety-critical systems. He has developed forensic guidance on behalf of the UK civil nuclear industry and helped develop European policy for the cyber-security of aviation – including ground based and airborne systems.