# Introducing Autonomous Systems into Operations: How the SMS has to Change

**John McDermid, Mike Parsons**

Assuring Autonomy International Programme (AAIP)

University of York, UK

**Abstract** *When an autonomous system is deployed into a specific environment there may be new safety risks introduced. These could include risks due to staff interacting with the new system in unsafe ways (e.g. getting too close), risks to infrastructure (e.g. collisions with maintenance equipment), and also risks to the environment (e.g. due to increased traffic flows). Hence changes must be made to the local Safety Management System (SMS) governing how the system is deployed, operated, maintained and disposed of within its operating context. This includes how the operators, maintainers, emergency services and accident investigators have to work to new practices and develop new skills. They may also require new approaches, tools and techniques to do their jobs. It is also noted that many autonomous systems (for example aerial drones or self-driving shuttles) may come with a generic product-based safety justification, comprising a safety case and operational information (e.g. manuals) that may need tailoring or adapting to each deployment environment. This adaptation may be done, in part, via the SMS. This paper focusses on these deployment and adaptation issues, highlighting changes to working processes and practices.*

## 2 Introduction

### *1.1 Background & Rationale*

Why a Safety Management System? Recent understanding of how accidents and incidents happen puts more emphasis on the causal factors external to the system and the organisational factors that contribute to errors being made (CAA, 2022).

The latter factors include how the organisation operates, how it sets out its procedures, how it trains its staff and what level of importance it gives to safety issues identified. A Safety Management System (SMS) addresses this and allows a proactive approach to safety by identifying causal factors and acting before an event happens. An SMS can therefore contribute to improving safety through a greater understanding of the hazards and risks affecting safety in the organisation.

In summary an SMS is the set of processes, procedures, management activities and cultural aspects that an organisation uses to ensure safety in its operation. Two useful definitions are:

> *"… a systematic and proactive approach for managing safety risks…[an] SMS includes goal setting, planning, and measuring performance. An effective safety management system is woven into the fabric of an organisation. It becomes part of the culture; the way people do their jobs*" (CAA, 2022), and:

> "*Safety Management Systems for product/service providers … integrate modern safety risk management and safety assurance concepts into repeatable, proactive systems. SMSs emphasize safety management as a fundamental business process to be considered in the same manner as other aspects of business management*." (FAA, 2022).

These two SMS definitions include a common set of four process areas for components of an organisation:

1. Safety policy and objectives (management commitment, plans, methods, processes, and organizational structure needed to meet safety goals);
2. Safety risk management (new or revised risk controls based on risk identification and assessment of acceptable risk);
3. Safety assurance (evaluates the effectiveness of risk control strategies; supports the identification of new hazards);
4. Safety promotion (training, communication, and other actions to create a positive safety culture).

Explicit SMSs exist in domains other than aviation. The European Union Agency for Railways (EUAR, 2022) has some concise statements that help to frame the nature of a typical SMS:

> *"The purpose of the SMS is to ensure that the organisation achieves its business objectives in a safe manner and complies with all of the safety obligations that apply to it…[it] enables the identification of hazards and the continuous management of risks related to an organisation's own activities, with the aim of preventing accidents…an SMS will provide an organisation with the necessary confidence that it controls and will continue to control all the risks associated with its activities, under all conditions…The SMS integrates into the business processes of the organisation…The SMS should be a living set of arrangements*

*which grows in maturity and develops as the organisation which it serves does so"*

This paper focuses on changes or additions to the SMS where autonomous systems (AS) are introduced into an existing environment.

## *1.2 Context*

We consider a new AS deployed into a specific environment. It could be an automated shuttle starting operations on a university campus; it could be an automated pallet system introduced into a factory; it could be a drone used by the military on the battlefield for the first time. It is assumed that this system will be delivered with a (product) safety case report, and that this comes with generic operational guidance and manuals covering a range of expected application situations. This guidance may reference a CONOPS (Concept of Operations) and also may have some operational restrictions or limitations which need to be observed. There may be a separate deployment safety case, demonstrating how the AS will meet safety claims for the specific environment, but this is unlikely for generic AS and is not assumed in the analysis presented here.

The AS is likely to be deployed in a staged process (see figure 1), involving[1]:

(i) Commissioning (pilot, trials, introduction into service),
(ii) Remote-controlled operations (possibly involving partial autonomy),
(iii) Fully autonomous operational service, and finally
(iv) Withdrawal from service.

In all these phases additional support is required via processes that include: (a) Monitoring, (b) Changes via maintenance and upgrades to functionality and (c) Incident and accident management. There may also be some generic supporting activities for the particular site or environment that the AS is working in; these include staff training and competency management.

[1] Stages could be based on the models of levels of autonomy used in, for example, the automotive sector. In this case, a vehicle may initially have an operator who can intervene while the vehicle drives itself before the later stages where there is no driver.

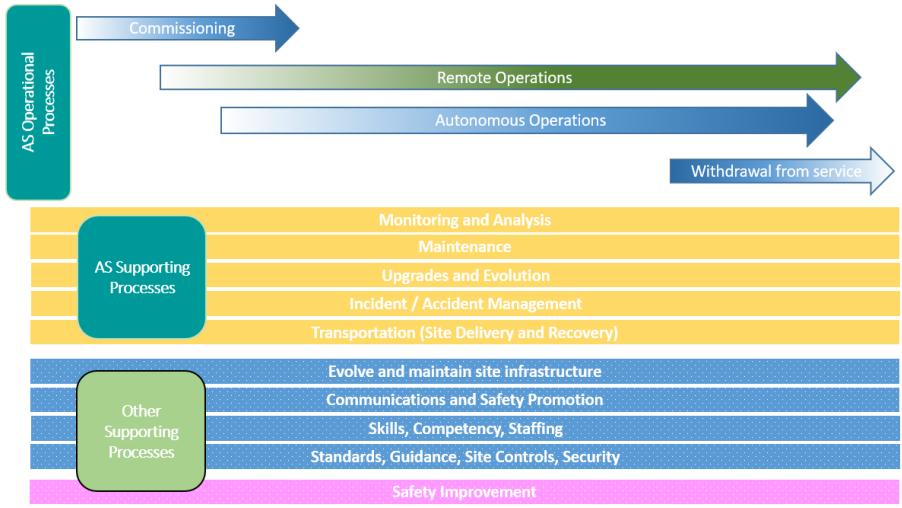


**Fig. 1.** Generic Deployment and SMS Areas for an Autonomous System

All these processes have to be updated for the introduction of an AS. For brevity, this paper focuses on five of these areas.

## *1.3 Structure of this paper*

The introduction in section 1 is followed by a description of the work currently being undertaken in the AAIP to produce a framework for production of additions to an SMS to support AS (informally the "SMS delta"). Section 3 is the main part of the paper that considers selected SMS areas and discusses the nature and type of the additions (and perhaps changes) required when introducing an AS. Section 4 presents some conclusions on the work so far, and section 5 outlines some areas considered for future work.

## 2 SODA

The University of York AAIP programme (AAIP, 2022) is currently developing a management framework for the Safety of Deployed Autonomous Systems, (SODA), (SODA, 2023) as part of a family of developments for AS including assuring the machine learning (ML) based system components (AMLAS, 2022) and assurance of an AS within a complex environment (SACE, 2022).


ASSURING
AUTONOMY
INTERNATIONAL PROGRAMME

Guidance on the Safety of Deployed Autonomous Systems (SODA)

John McDermid, Mike Parsons, Richard Hawkins and Matt Osborne

Assuring Autonomy International Programme (AAIP)
University of York

Version 0.1, July 2022

The material in this document is provided as guidance only. No responsibility for loss occasioned to any person acting or refraining from action as a result of this material or any comments made can be accepted by the authors or The University of York.

This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA. Requests for permission for wider use or dissemination should be made to the authors.

Contact : firstname.lastname@york.ac.uk.

SODA produces the AS-specific elements of the SMS for operation of the AS at a specific site. SODA is a process for systematic construction of an "SMS delta", i.e. the changes required to the SMS to enable safe operation of an AS. It assumes that there is already an SMS in place for operations at the site.

The result of applying SODA is a set of AS-specific processes and procedures to add into a standard SMS, including identifying a set of tangible inputs and outputs (documents and other artefacts).

SODA comprises a set of processes covering the activities identified in Figure 1. Each process comprises a set of process steps with inputs and outputs for each step. The process for the Commissioning activity from Figure 1 is shown in Figure 2. By working through the process steps, an addition to the SMS will be produced supporting the pre-operational commissioning of the AS at that site. The first three stages produce procedure fragments that are drawn together into a coherent procedure based on the supplied template.

The intent is that the SODA processes will be undertaken by safety professionals – site safety managers, safety engineers, etc. with the resultant processes and procedures produced in company-standard form to inform operators on the site or in a remote operations centre (ROC).

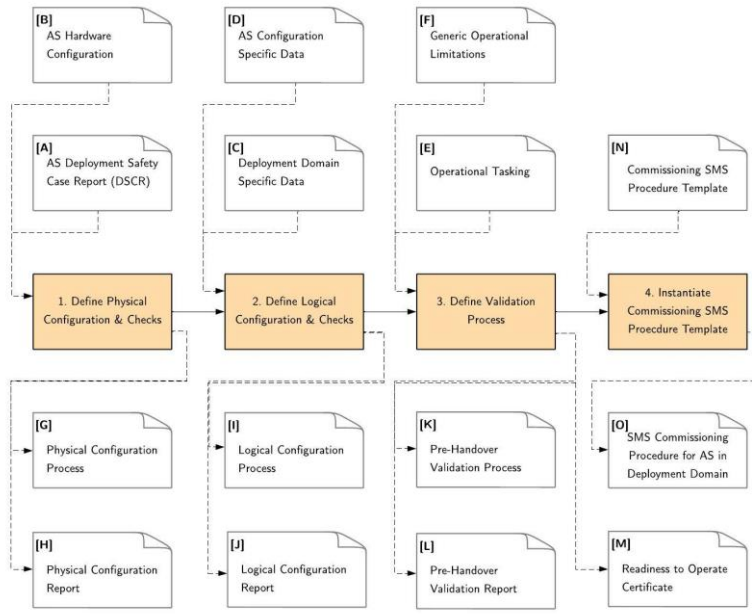


**Fig. 2.** Illustrative SODA Process for Commissioning

## 3 Selected Additions

This section considers some of the additions to the SMS needed for safe deployment of an AS, as might be expected to be derived from SODA. There could be many areas of the SMS that need updating. In this paper five of these areas are examined:

1. Remote Control and Autonomous Operations
2. Monitoring and Analysis
3. Upgrades and Evolution
4. Skills, Competency, Staffing
5. Incident / Accident Management

Figure 2 below shows how these areas relate to the original SMS scoping:

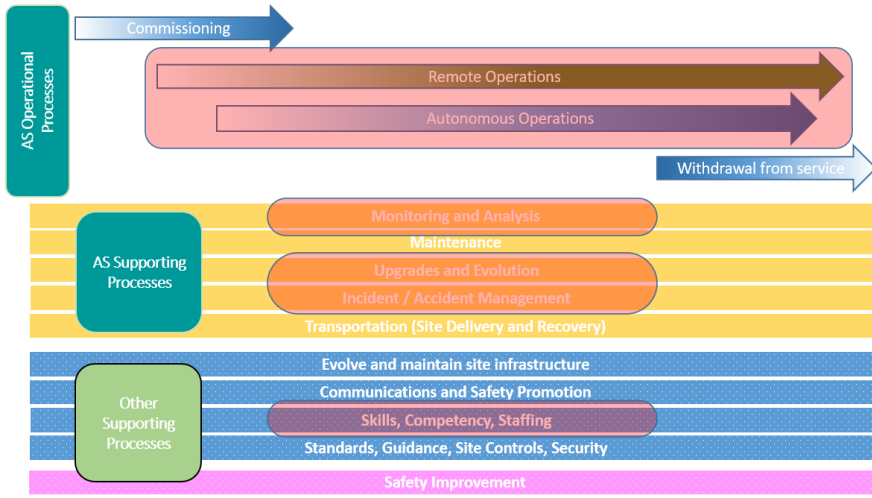


**Fig. 3.** Coverage of selected SMS areas from Figure 1 in this paper

So, in each of the four typical SMS areas we consider the following as needing review and/or update:

1. **Safety policy and objectives**. This needs to include contracts and agreements with the AS supplier and maintainer including Service-Level Agreements (SLA), that may also need to involve the site operator. These will include things like hours of support and call-out times. They will also include agreements - how much the AS supplier will attempt to do remotely and how often they may be required to attend site. Objectives should also be updated to cover the autonomous aspects, given that certain aspects may not be predictable such as object recognition via machine learning. Targets for safety performance of the AS in its context should be set. Policies that may need to be in place regarding the AS include communications recording and playback facilities, data storage and retention, data sharing protocols and engagement agreements with certification / accreditation bodies, regulators[2] and accident investigators. It is important to ensure that everything is in place so that any accidents or incidents involving the control of the AS can be thoroughly investigated using replayed data.
2. **Safety risk management**. This SMS element needs to consider issues of loss of communications, remote/local/autonomous behavioural conflicts, and interactions with personnel on site, plus other autonomous systems, etc. New pro-

[2] It is important to note that in certain sectors (e.g. nuclear, aviation) the regulator has an important role, and the agreements and methods of engagement put in place may need to be comprehensive. For instance, a regulator in these industries may want to be involved in reviewing any change applied to the AS, before it goes live.

cesses and protocols may need to be defined - on-site maintenance where remote operation must be disabled and autonomy isolated. New hazards relating to conflicts may be added by personnel who have to interact with the AS on a daily basis.

3. **Safety assurance**. The frequency and type of monitoring needs to be defined so that all issues can be recorded, analysed and eventually their resolutions audited. The effectiveness of the SMS on the organisational safety performance must be evaluated and reviewed. When issues are found there may need to be temporary changes to operations (e.g. a fault found with the autonomy means that the AS is disallowed in certain areas of the site). Clearly the existing Safety Case for the site(s) will need to be reviewed in the context of including the AS.
4. **Safety promotion**. Staff will need to be trained and regularly informed about how to interact with the AS, and what functionality is currently enabled. They will need a mental model of the AS behaviour that can be developed and enhanced through regular briefs. They need to be confident that the AS is fully under local control (i.e. autonomy isolated) when performing various duties, e.g. recharging and cleaning.

## *3.1 Generic Additions to the SMS*

Generic additions to the SMS will be required taking into account the following (some of which can be seen as inputs to the Commissioning process in Figure 2):

**Table 1.** Generic Additions to the SMS

| Name | Description |
|---|---|
| Operating Scenarios | Scenarios (typically defined as Use cases) that describe the expected behaviour patterns of the AS at the site. |
| Deployment Domain Specific Data | This may include maps of the deployment domain indicating key points, e.g. ingress and egress, no-go areas for the AS, etc. Local information may be needed, e.g. areas of the site prone to communications dropouts, the location of the "home" charging station, and "muster points" in the event of emergencies. |
| Generic Operational Limitations | Constraints on the operational design model (ODM) that will apply to the AS for an extended period (perhaps throughout its operation), e.g. temporal (can't operate outside working hours), or geographical (terrain where the AS is likely to get stuck, and thus must avoid), etc. |
| Operational Tasking | The task or set of tasks that the AS is expected to do, e.g. traverse the site looking for particular features or performing maintenance actions. |

| Name | Description |
|---|---|
| Hazardous Scenarios | A subset of the scenarios that are identified as hazardous; note that these will likely be included in the Deployment Safety Case report (if there is one). |
| Existing Site SMS / Safety Procedures for Site | The pre-existing safety management system for the site(s). This may consist of higher-level documents that flow down to all sites. This may be specific to the contractor(s) that operate and maintain the site. |

## *3.2 Remote Control and Autonomous Operations*

When an AS is deployed it is often able to be remotely controlled as well as capable of fully autonomous operation[3]. However, this ability creates new risks and complexities as these modes of operation may conflict. An example might be if the AS is a vehicle, and the remote operator mistakenly drives the AS at another vehicle or a static object potentially causing a collision. In this case the AS may try to avoid the collision by disobeying the remote-control commands.

Hence, there may need to be operational rules (protocols, procedures, check lists) added to the SMS to ensure that a remote operator is fully aware of the situation the AS is in before issuing commands.

In reality, the interaction between the operator and the AS is likely more complex than this: the AS may have layered levels of functionality (including avoidance of harm behaviours and self-protection), and the remote control may only be able to override some of the functionality, even if trying to avoid an accident.

Who has control of the AS and who is responsible when accidents occur could be a very difficult problem with typical remote, local and autonomous control: there may be several levels of remote operation, e.g. via an operations centre and via a local hand-held controller. Conflicts with both these and the autonomous functions need to be managed through SMS procedures and protocols, at least in part.

In this context, the SMS will need to take account of:

[3] In some industries (e.g. UK civil aviation) autonomous operation is defined as **only** where there is no possibility of human intervention. If there is a remote link in place and a remote operator can intervene (to some extent) then it is not autonomous but rather "automatic with high authority", and this is an important distinction. In some safety-critical industries, SMSs have been developed over many years to deal with those types of system and interactions involved. The real issue is where intervention is not (realistically) possible, e.g. some vehicle scenarios (land/air) or a cobot that works in a factory in an uncaged environment alongside humans. Also where AI is used to support the AS decision making; here SMSs need to be updated to account for non-deterministic (or at least not easily explainable) behaviours and decisions made and their effect on staff.

**Table 2.** Remote Control and Autonomous Operations

| Name | Description |
|---|---|
| AS Message Flow Definition | Messages sent to and received from the AS either from a remote operations centre (ROC) or a hand-held device. |
| Primitive Procedures Definition | The set of basic procedures for interacting with the AS, e.g. to isolate autonomy (and report that this has been done), to start a pre-stored task, to report completion of the task, to "request assistance", etc. |
| Remote Operations Centre Procedures | Processes and procedures for the operations staff to follow at the ROC (may cover many ASs and many sites). |
| Local Operating Protocols | Step-by-step instructions as to how to manage local operation and autonomous operation of the AS at site |

## *3.3 Monitoring and Analysis*

In conjunction with the local / remote control issue, there is likely to be a need to have real-time or near real-time monitoring of an AS, to ensure everything is working within safe bounds. This is most likely done using radio networks that are subject to drop-outs, delays and interference (both unintended and intentional). Hence processes and procedures in the SMS are required to establish what to do if communication is lost. Typically, a certain level of communication loss is tolerated but after a period must be re-established or the AS will have to perform contingency actions, depending on the circumstances (e.g. execute some sort of minimum risk manoeuvre or abort its mission)[4] and enter a state enabling it to be recovered.

Analysis is required of the safety performance of the AS. This could be near real-time (e.g. via a safety dashboard) or slower-time analysis (each day, week or month). This requires that data is available from the AS, so must either be transmitted or stored for later uploading. Potentially, the amount of data produced will be very large, and it will be a significant effort to process it to look for early signs of faults that might lead to a safety event, or indeed no-fault cases that might cause a problem[5]. Processes and procedures for collecting, storing, processing and analysing the data from the AS will therefore be required in the SMS. These will likely have to involve third-party organisations to provide communications services, plus possibly cloud storage provided by web service providers.

[4] It is recognised that for aerial AS it may be safer to continue to an intermediate safe location or indeed to complete the mission if it is close to a landing zone or airport

[5] In this case although an individual or fleet of AS are working to their specification, the combination of behaviours (perhaps involving other manufacturers' AS or interactions with humans) is leading to unsafe situations

**Table 3.** Monitoring and Analysis

| Name | Description |
|---|---|
| AS Monitoring Data | Data from interoception, assessing the state of the hardware, e.g. from actuator built-in tests (BIT), and from assessment of the impact of the environment on the sensing suite, e.g. impairment due to fog or rain. This might inform temporary operational limitations. |
| Environmental Monitoring Data | Information about the state of the deployment including from sensors within the infrastructure and potentially from the AS itself or other AS on site. |
| AS Safety Performance Analysis | A report produced regularly that demonstrates the required safety performance is being met. Compilation of this report may require data from the manufacturer as well as from site. |
| Agreements with Communications Providers | Increased site communications needs may require changed agreements with communications providers, covering specific service levels including assurance and integrity targets |
| Agreements with Cloud Storage Providers | There will be a large amount of data produced by the AS and associated infrastructure and this will need to be stored in a secure cloud environment. |
| Data Sharing and Retention Policy | It is critical that data that has been saved (either stored in the AS or via site infrastructure) and is not lost (either deleted or overwritten). It must also be able to be shared with the appropriate parties: manufacturer, maintainer, site owner, independent investigator, etc[6]. |

## *3.4 Upgrades and Evolution*

Changes to an AS will typically be made via Over-The-Air (OTA) updates that can be done on site or in the field wherever communications are possible. OTA updates will be governed by the SMS (for instance, there may be restrictions on when an update can become active, and where it can be trialled) that may affect both the operating software and data used within the AS and site infrastructure. The data used by the AS may be of several different types, including:

1. Configuration Data (to configure features within the AS itself)
2. Navigation Data (including maps, allowable routes, prohibited areas, etc)
3. Site Data (including changes to site infrastructure locations, etc)

[6] It is recognised that some data may need to be post-processed (e.g. anonymised) before sharing, for example, to obfuscate faces of people that a camera may have captured. GDPR requirements may apply.

4. Machine Learning Training Data, that may influence behaviours such as navigation or object recognition

What is different here is that the software and data will largely govern the behaviour of the AS, and that the changes in behaviour need to be understood by people working on site.

Of course, changes may also be introduced to the hardware of the AS. Sensors could be replaced, functionality could be upgraded (e.g. higher-capacity batteries) and new features added (e.g. additional cameras). Changes may be undertaken at site or the AS may have to be returned to the manufacturer for the change to be performed. Different SMS processes and procedures are required in each case, for instance, AS removal from site and AS delivery to site. In all cases the disabling of autonomous functions, and verification of this action is paramount and procedures for this will have to be built into all SMS processes and supported by the AS itself. There may also need to be a proving area or testing ground at site where changes can be tested and verified, again this would largely come under site procedures within the SMS[7].

**Table 4.** Upgrades and Evolution

| Name | Description |
|---|---|
| Change Management Procedures | The various procedures and processes to be used to make changes to the AS and the supporting infrastructure. May involve change at different locations and different types of change (hardware, software, data). |
| OTA Update Protocol | The steps to enable safe changes via OTA updates. Should cover fleet upgrades, mixed fleet issues and backing out unwanted changes. |
| Upgrade at Site Procedures | Procedures for making change at site, including isolation of autonomy. |
| Testing Ground Definition | If the changes are to be tested at site then a definition of the test ground (environment) will be needed. |
| Testing Ground Procedures | If the changes are to be tested at site, then detailed procedures will need to be established for testing in a safe manner, away from operational infrastructure. |

## *3.5 Skills, Competency, Staffing*

The SMS will have to cover areas of staff training and competency for dealing with the AS. Depending on the nature of AS, this could range from simple awareness

[7] It is recognised that there may be local site operators operating under service contracts with their own SMS in place. In this case the prime should ensure that any higher-level SMS requirements related to the AS flow down to local operators as needed.

courses (for an AS unlikely to cause any harm) to specific and detailed training with examination and certification for larger, faster or perhaps airborne AS. This training will likely have to be tailored for the specific site to include local conditions (for instance, including procedures to deal with the muddy conditions or flooding for a land vehicle AS). The training should be such that it enables staff that may come into contact with the AS to interact with it safely and to minimise any operational difficulties for them or the AS. The SMS will have to contain mechanisms for staff to report issues with the AS, and to be able to be informed of updates and changes in behaviour as a result of reports.

Of course, staff do not always behave as they should, and it is possible that protective actions are forgotten or ignored (e.g. not isolating autonomy before maintenance actions). Also, unauthorised actions or dangerous interactions with the AS may take place (for instance, 'playing chicken' in front of a vehicle AS), putting people at risk. In this case the SMS needs to anticipate as much foreseeable misuse as possible and contain warnings in manuals and provide regular training covering misuse. It may have links to site disciplinary procedures to deal appropriately with any actual misuse to discourage recurrence.

**Table 5.** Staff Training & Safety Information

| Name | Description |
|---|---|
| Staff Handbook | The staff handbook should outline how staff are expected to work with the AS, and detail warnings and limitations. |
| Staff Training Courses | Staff may require training before being allowed near the AS. |
| Staff Briefings | Staff may require regular briefings if the AS functionality changes due to frequent OTA updates. |
| Incident and Fault Reporting | There may have to be changes to the standard site incident reporting procedures (e.g. additional statutory information required) when logging an incident involving an AS. |
| Staff Welfare and Support | Existing site staff may well feel threatened by the AS if it performs duties previously done by them. They may need retraining, redeployment and support services to manage the safe introduction of the AS. |

## *3.6 Incident / Accident Management*

An operational AS will require processes and procedures for managing accidents and incidents; dealing with them is an important part of any SMS. This will include everything from managing communications with the press and accident investigators, through analysis of the causes, to making the accident site safe.

Hence, with an AS, the parts of the SMS that require modification include policies for communications (and site security[8]), noting that there may be a lot of press interest in a major accident involving an AS, changes to procedures such as those used for accident management including how to establish a safe site, instructions on how to verify that the autonomous functionality is disabled (and cannot be mistakenly re-enabled remotely), and how to recover the AS. There may also be additional processes for how to deal with investigations internal to the organisation and also with external accident investigators. If people are involved in the accident there may be injuries to deal with and emergency services may need to operate special protocols when autonomy is involved, requiring independent verification of autonomy isolation to ensure that emergency services staff are not put at undue risk.

Hence the existing SMS may need to have updates related to:

**Table 6.** Accidents & Investigations

| **Name** | **Description** |
|---|---|
| Agreements with Manufacturers and Maintainers | It is important that the manufacturer of the AS and any maintenance organisation are 'on side' and able and willing share information and assist with any incident recovery and analysis. |
| Agreements with Independent Investigators | Agreements, permissions and working methods need to be in place with any independent investigators in advance. |
| AS Manufacturer Supporting Information | Any additional information that the manufacturer of the AS has regarding management of incidents or accidents, e.g. autonomy isolation, towing considerations for land-based AS, etc. |
| Site Information for Accidents | Any site-specific information regarding accident management, e.g. site procedures for turning off power; isolation of autonomy; chain of command; emergency services call-outs; fire routes, etc. |
| Tools and Equipment Required | The AS may require special handling, tools and equipment (e.g. for recovery and towing). |
| Skills and Competencies Required | The training, skills and competencies staff must have in order to deal with an accident involving the AS. |
| Incident Handling Process Definition | A site-specific process defining the staff, procedures and actions relating to handling a site incident involving the AS. May involve everything from handling the press to isolating part of the site. |

[8] It is recognised that the inter-relationship between an SMS and security management processes is an important one. There can be issues and conflicts to resolve (for instance in sharing of operational data about the AS and access to certain areas of a site for accident investigation). This will be the subject of future work.

| Name | Description |
|---|---|
| Configuration, Software and Data State of the AS | It is important that the configuration data and software state of the AS is preserved after an incident so that it can be analysed[9]. This is likely to require a detailed procedure on site, especially if communications with the AS is lost. |
| Configuration, Software and Data State of Site Infrastructure | It is important that the configuration data and software state of the site infrastructure (including other AS) is preserved after an incident so that it can be analysed. This is likely to require a detailed procedure on site. |
| Temporary Operational Limitations | Constraints that apply for a limited period of time, e.g. spatial limitations whilst recovery operations are carried out, or temporal limitations due to what is happening on site. |

## 4 Conclusions and Discussion

This work has examined changes to an existing SMS when an AS is introduced. Five areas were discussed, and changes proposed. The nature, scope and scale of changes will depend on the deployment context (e.g. land, air, water, space) and characteristics of the AS (e.g. size, weight, actuators, proximity to staff). However, the generic AAIP SODA framework is designed to be comprehensive and detailed enough to be applicable to a wide range of AS modalities. Whilst not discussed in detail above, it is expected that developing the SMS will identify derived requirements on the AS, e.g. to be able to report that autonomy is isolated, to move to a "muster point" if informed of a site emergency. Thus, a level of co-design between the AS and the SMS may be valuable

It is acknowledged that an SMS is rarely developed from scratch, instead it is a combination of knowledge, process, procedures and instructions that evolves over time and generally develops incrementally. Therefore, it has been assumed that only changes (deltas) need to be made to cover the AS introduction. However, it may not be so simple, as there may be conflicts with existing processes and procedures, and these may also need to be updated. Lastly, it may not be obvious what should be done in cases not covered by the SMS (perhaps collisions with other diverse AS that should not be on site, or unexpected human behaviours when interacting with the AS). Therefore, it is recommended that a full review of the SMS is undertaken after introducing the changes for the introduction of an AS.

[9] Note that local storage of data (vehicle, telemetry) for accident investigation purposes may impose design requirements on the AS, necessitating a comprehensive "black box" in effect; the SMS may also need to support data transfer and storage as well. This is the subject of future work.

## 5 Further Considerations

This section outlines some additional work that could be undertaken to further elaborate what is required for an SMS dealing with operations incorporating an AS.

Firstly, the other identified areas of the SMS require analysis to see what changes may be required. In addition, the following areas are likely to be important considerations when developing the SMS:

**Regulators:** In many safety-related industries (at least in the UK) there is strong regulation in place, backed up by legislation. This applies particularly to high-risk and well-established sectors such as nuclear, civil aviation and rail. It is expected that the regulators in these industries will take an active role in the introduction of an AS, setting out objectives and requirements, producing guidance, reviewing and approving changes to an AS, monitoring operational safety performance, and making recommendations for safety improvement. Regulators may want to be involved in monitoring other aspects such as staffing issues (effects on absences, rostering, etc.) and the impact of introducing AS on safety culture within an organisation.

**Digital Twins:** When an AS is deployed into an operational environment there may be a need to integrate it into the maintenance and operational models that the organisation uses to monitor and maintain the operational status. For complex sites the models utilised can be very sophisticated. For this purpose, it may be necessary to create a digital twin model of the AS in its environment. This model can be as abstract or detailed as required but may have to include modelling of autonomous decision making[10]. As an example, if the autonomous decision making within a land vehicle (say farm vehicle) always chooses a particular route over rough ground, this ground may become muddy and impassable and therefore increase risk over time.

**Replacement of the AS:** When an AS is upgraded for a newer model or different variant, the SMS will need to be reviewed to see if any further changes are required. In fleet situations, the picture is more complex as generally not all AS can be replaced in one go and operation with a mixed fleet of, say, older and newer models together may be more likely. In this case the risks of operating with different AS models or versions together must be assessed and mitigated.

**Removal of the AS:** Where an AS performs a safety function, e.g. fire detection or suppression, then it cannot be removed without an increase in risk. This risk will have to be managed, either manually (e.g. increased monitoring by staff), or by replacing the AS by other systems (that may or may not have autonomous functions).

[10] Some digital twin models can take data from the physical system to continually refine the digital model.

**Phased Introduction:** It is recognised that an AS is usually not introduced into an operational setting in one go; typically, there may be a series of iterative trials, phases or stages where more functionality is exercised and additional parts of the domain (say areas of the land site or air space) are included in each phase. In this case, full safety documentation is unlikely to be available at the start of trials, so it is important that appropriate mitigations are put in place to address the risks at each stage, noting that these might be progressively reduced as confidence is built in the safe operation of the AS.

**Acknowledgments** We thank Paul Hampton for very valuable review comments.

**Disclaimers** All views are those of the authors and not their respective organisations.

**References**

AAIP (2022), Assuring Autonomy International Programme, https://www.york.ac.uk/assuring-autonomy/ , accessed October 2022

AMLAS (2022), Assurance of Machine Learning for use in Autonomous Systems (AMLAS), AAIP, https://www.york.ac.uk/assuring-autonomy/guidance/amlas/ accessed November 2022

CAA (2022), Safety Management Systems: Guidance for small, non-complex organisations, CAP 1059, https://publicapps.caa.co.uk/docs/33/CAP%201059%20SMS%20for%20small%20organisations%20(p).pdf , accessed October 2022.

CAA (2022a), Safety Management Systems (SMS): guidance for organisations, CAP 795 https://publicapps.caa.co.uk/docs/33/CAP795_SMS_guidance_to_organisations.pdf, accessed October 2022

EUAR (2022), ERA->Activities->Safety Management System, https://www.era.europa.eu/activities/safety-management-system_en , accessed October 2022

FAA (2022), SMS Explained, https://www.faa.gov/about/initiatives/sms/explained , accessed October 2022

SACE (2022), Guidance on the Safety Assurance of autonomous systems in Complex Environments (SACE), https://www.york.ac.uk/assuring-autonomy/guidance/sace/ , accessed November 2022

SODA (2023), Guidance for the Safety of Deployed Autonomous Systems (SODA), AAIP University of York, in preparation