TECHNICAL REPORT

ISO/TR 9241-810

First edition
2020-08

Ergonomics of human-system interaction —

Part 810:
Robotic, intelligent and autonomous systems

Reference number
ISO/TR 9241-810:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 159, *Ergonomics*, Subcommittee SC 4, *Ergonomics of human-system interaction*.

A list of all parts in the ISO 9241 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Product development of systems with robot, intelligent and autonomous characteristics is rapidly progressing. Given the human-system issues of such systems, timely guidance covering these issues is necessary to help all sectors of industry to design, field and operate first-time quality robotic, intelligent, autonomous (RIA) systems, and build appropriate trust in products and services that use these systems.

There is an urgent need for a Technical Report from ISO explaining the existing, emerging and potential human-system issues and consequences for use and users associated with systems that have robot, intelligent and autonomous characteristics. This document explains the existing, emerging and potential human-system issues and consequences for use and users associated with systems that have RIA characteristics. It identifies the potential risks and priorities for standardization to address these issues. Solutions will be the subject of future standards.

This document reviews the ergonomics for a range of RIA systems. It describes the human-system issues that should be considered in the application of these technologies and identification of priorities for future standardization work. The purpose of this study is to identify and explore the ramifications of a categories of issues involving RIA systems that suggest a need to reset the boundaries of what is called ergonomics. The conclusion is that to make an ergonomic RIA system, the practice of ergonomics will need to do more, working together with new disciplines, and can require new tools, methods and approaches to support the design and integration of these types of systems into working environments and organizations. Ergonomics will also need to identify relevant research from a wide variety of scientific disciplines, as well as conducting our own research to ensure we have a robust evidence base to guide the development of these systems.

The paradigm behind human-systems interaction standards so far has been that of tool use. The ISO 9241 series is for interactive tools and the physical environment within which they are used. RIA systems necessitate a new paradigm. Agents developed using these technologies will be more connected, complex, probabilistic and non-deterministic, social, and augment human capabilities well beyond merely replacing physical work. Interaction with these agents can become a relationship, their interface a personality, and users and agents can form complex human-machine teams, working together towards a shared goal.

The evolution of RIA systems will significantly alter the nature of tasks users perform. The design of work will likewise be altered. Applications of RIA systems represent a significantly more complete and impactful replacement of human activity than has been seen with any other form of technological labour-saving device. For example, when working with another person on a common task, how do you diagnose a failure state in your interactions? How are you to interpret the off-nominal behaviour of a team member? How are you to interpret and predict the behaviour of other people who are operating within the same environment as you are but are otherwise not directly coordinating activity? What is the safe state you can fall back on in the event of a failure in your interaction with another person? Now, replace that person or team member with an RIA system. The changes in the nature of tasks and the design of work to accommodate the complex, social human-machine interaction of an RIA system is fundamental for ergonomics, but will require that the ergonomics community adapt its best practices and expand into areas of psychology and sociology that few ergonomists deal with on a regular basis.

The focus of this document is breadth not depth, and issues not answers. The emphasis is on describing general issues and the consequences of not addressing them, even though not all will/can be relevant to all types or applications of RIA systems covered by this document. But be sure that this is the case for your application, and that you take account of the categories of issue and context that do apply.

Ergonomics of human-system interaction —

Part 810:
Robotic, intelligent and autonomous systems

1 Scope

This document addresses:

— physically embodied RIA systems, such as robots and autonomous vehicles with which users will physically interact;

— systems embedded within the physical environment with which users do not consciously interact, but which collect data and/or modify the environment within which people live or work such as smart building and, mood-detection;

— intelligent software tools and agents with which users actively interact through some form of user interface;

— intelligent software agents which act without active user input to modify or tailor the systems to the user's behaviour, task or some other purpose, including providing context specific content/information, tailoring adverts to a user based on information about them, user interfaces that adapt to the cognitive or physiological state, "ambient intelligence";

— the effect on users resulting from the combined interaction of several RIA systems such as conflicting behaviours between the RIA systems under the same circumstances;

— the complex system-of-systems and sociotechnical impacts of the use of RIA systems, particularly on society and government.

This document is not an exploration of the philosophical, ethical or political issues surrounding robotics, artificial intelligence, machine learning, and intelligent machines or environments. For matters of ethics and political issues, see standards such as BS 8611 and IEC P7000. However, this document does identify where and why ethical issues need to be taken into account for a wide range of systems and contexts, and as such it provides information relevant to the broader debate regarding RIA systems.

This document has a broader focus than much of the early work on autonomy that relates to the automation of control tasks and mechanization of repetitive physical or cognitive tasks, and centres on levels of automation.

Although this document addresses a wide range of technology applications, and sector and stakeholder views on the issues, the treatment of each can be incomplete due to the diverse and increasingly varied applications of RIA systems.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

No terms and definitions are listed in this document.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

4 Symbols and abbreviated terms

AI	artificial intelligence
CRM	crew resource management
DM	decision making
GPS	global positioning system
HCD	human-centred design
HCI	human-computer interaction
HCQ	human-centred quality (see ISO 9241-220)
HF	human factors
IA	intelligent agent
ICT	information and communications technology
IVR	interactive voice response
ML	machine learning
OODA	observe–orient–decide–act
RIA	robotic, intelligent, autonomous
RPA	robotic process automation
UxV	unmanned (where x = space, air, ground, surface, sub-surface) vehicle
UI	user interface
UX	user experience

5 Report contents and structure

The target audience for this document is decision-makers, designers and engineers who would benefit from the consideration of human-systems issues of RIA systems. Futurists, researchers, technology developers, regulators and legislators can also find this document useful.

The target audience for this document is the standards development community and ergonomists involved in developing, acquiring and/or commissioning RIA systems.

This document is based on an analysis that projects forwards from current applications of technology to more connected, complex, probabilistic and non-deterministic, social systems/entities/agents, and human augmentation. Social in this context also includes physical interaction. Applications considered include robots, intelligent systems and environments such as smart buildings that control or otherwise influence an environment, and autonomous agents/systems. The analysis considers views and concerns of: RIA system users and stakeholders from various industry sectors regarding the impact on future job roles, human tasks and organizational structures, safety, system trust, rights and culture. The limits for ergonomics are considered together with an initial identification of potential areas of change. A broad range of published sources and expertise was drawn on during the creation of this document. It includes the futurology literature, regulatory work, input from astute observers and reports of current and planned products. Extensive discussion and analysis by the project team is also included.

— Clause 6 discusses relevant concepts in AI and ergonomics.

— Clause 7 describes the groups of identified issues.

— Clause 8 describes the hazards and possible harm that can result if Ergonomics is not applied.

— Clause 9 describing how various existing ergonomics standards address the issues.

— Clause 10 describes the changes in ergonomics standards required to better address RIA systems technology.

Annexes A to E are written for:

— the ergonomics community — to give their input to RIA system projects/discussions face validity, provide food for thought regarding how ergonomics can be applied/should evolve/needs to be supported, gives a framework for issues to raise if involved with such projects;

— those developing, acquiring, commissioning or approving RIA systems — providing a set of considerations and potential issues to think about for those in any executive, project, design or legal and regulatory role;

— developers and users of standards who need to understand how the ergonomics aspects of RIA systems affect their activities — alerting those who have not so far included human or ergonomic requirements in relation to RIA systems in their domains to new or emergent human-system issues or needs.

Annex A elaborates the human-system issues within each category. Annex B presents examples of RIA systems, illustrating the issues, hazards, and ergonomics considerations. Annex C provides a two-stage review of the areas in which ergonomics needs to develop to address these issues. Annex D contains a more detailed description of the analysis and notes on the necessary extensions to ergonomics and standards. Annex E describes the analysis on which this document is based.

6 Concepts

6.1 General

There are many technologies used to implement RIA systems, various combinations of which are employed across a huge range of applications with which humans will interact. This has led to a general lack of agreement and precision in definitions and terminology, including those within standards where RIA system technologies and applications are defined in various ways according to specific requirements of the given context. As it is not possible to fully predict the different ways in which such technologies will be developed and applied in the future, this document does not refer to existing definitions from other standards. Instead, this document uses generic and commonly used terms because, although these can still invoke different individual interpretations and opinions, they are more generally and widely understood.

This document uses the most common generic terms in the title (robotic, intelligent, autonomous) with the understanding that they can trigger a range of associations and differences of opinion. These are not conceptually independent. Furthermore, this document focusses on their use by humans as collective descriptions for characteristics of types of intelligent agent. These agents are often qualified as to type or context of use (for example, autonomous car, intelligent building, care robot).

6.2 IT concepts

6.2.1 Intelligent agent

An intelligent agent is an entity which observes through sensors and acts on an environment using actuators. It directs its activity towards achieving goals. Intelligent agents can learn or use knowledge to achieve these goals. They can be very simple or very complex.

Intelligent agents are often software entities that carry out some set of operations on behalf of a user or another program with some degree of independence or autonomy, and in so doing, employ some knowledge or representation of the user's goals or desires. For example, autonomous programs used for operator assistance or data mining (sometimes referred to as bots) are also called "intelligent agents".

6.2.2 Autonomous agent

An autonomous agent is an intelligent agent operating on an owner's behalf with a high degree of independence.

Such an agent is a system situated in, and part of, a technical or natural environment, that senses any or some status of that environment, and acts on it in pursuit of its own agenda. The agenda evolves from drives (or programmed goals). The agent acts to change part of the environment or of its status and influences what it sensed.

Non-biological examples include intelligent agents, autonomous robots and various software agents, including artificial life agents, and many computer viruses. Biological examples are not yet defined (apart from living organisms).

NOTE Autonomy is a system property; it does not necessarily imply artificial intelligence.

The term machine learning is often used in conjunction with intelligent agents and some definitions of an autonomous system include the ability to learn as a characteristic of such systems.

6.2.3 Machine learning

Machine learning (ML) is a field of artificial intelligence that uses statistical techniques to give computer systems the ability to "learn" (e.g. progressively improve performance on a specific task) from data, without being explicitly programmed.

6.2.4 Autonomous robot

An autonomous robot is a robot that performs behaviours or tasks with a high degree of independence. This feature is particularly desirable in fields such as spaceflight, household maintenance (such as cleaning), waste water treatment and delivering goods and services.

Some modern factory robots are autonomous within the strict confines of their direct environment. It may not be that every degree of freedom exists in their surrounding environment, but the factory robot's workplace is challenging and can often contain chaotic, unpredicted variables. The exact orientation and position of the next object of work and (in the more advanced factories) even the type of object and the required task is determined. This can vary unpredictably (at least from the robot's point of view).

One important area of robotics research is to enable the robot to cope with its environment whether this is on land, underwater, in the air, underground or in space.

NOTE An autonomous robot is an embodied intelligent agent.

A fully autonomous robot can:

— gain information about the environment;

— work for an extended period without human intervention;

— move either all or part of itself throughout its operating environment without human assistance;

— avoid situations that are harmful to people, property or itself unless those are part of its design specifications;

— an autonomous robot can also learn or gain new knowledge like adjusting for new methods of accomplishing its tasks or adapting to changing surroundings;

— like other machines, autonomous robots can require regular maintenance.

6.2.5 ISO robot

The use of the term “robot” in this document is intended to include devices covered by the definition provided in ISO 8373:2012, 2.6.

The broader use of this term in this document is intended to accommodate variations and overlaps in the conceptualization, design and application of robots/robotics, and to avoid more specific definitions (like the definition in ISO 8373) that can pertain only to individual models and applications of robot as capabilities continue to evolve.

6.3 Ergonomics concepts

6.3.1 Ergonomics concern for RIA systems

There is a diverse range of interacting influences to be considered in the design of work involving RIA systems. Adaptation/evolution of ergonomics and its standards are required to better address RIA system issues/problems. Some of these topics are presently not within the scope of ergonomics to address. Ergonomics has traditionally concerned itself with physical and cognitive acts in a physical world. It is now addressing interactive systems and the digital world that extends to information, both about the physical world and of itself including science, business information systems, entertainment and knowledge. With RIA technology, ergonomics can also need to consider social ergonomics in order to assess and assist user experience, accessibility, usability and avoidance of harm. Ergonomics can need to extend its scope to take account of the effect of social interaction with and through RIA systems including social media, human-machine teaming, and adaptive interfaces and environments. Likewise, ergonomics can need to explore the impact to non-users of RIA systems who are in the environment in which the RIA systems are operating but are not using the RIA systems. Ergonomics can also need to take account of legal, regulatory and governance requirements that are being set by other parties, largely or even completely without ergonomics input.

6.3.2 Design approaches for RIA systems

There is a range of design approaches that can be taken when designing the relationship between an RIA system and its users, both human and other RIA systems. These are described in Table 1. They have ergonomic, sociotechnical, ethical and possibly political implications. Ergonomics applied to RIA systems should consider the most suitable paradigm for an application and the effect on individuals and society.

Table 1 — Design approaches for RIA systems

Design approach	Description	Benefit/Caution
Augmentation	The system improves human performance (including decision aids, exoskeletons, physical and/or cognitive prosthetics).	Extends human performance and capability. Possible health and ethical issues regarding elective use of augmentation for some advantage either to individual or society.

Table 1 *(continued)*

Design approach	Description	Benefit/Caution
Replacement	The system replaces human functions and/or entire human jobs.	Good for disliked, hazardous, repetitive or mundane jobs and those where there are skill shortages in the labour pool and where productivity can be improved. Unthinking application of replacement for specious purposes (such as unsupported safety and financial benefits).
Remoting	Allows the user to act on the physical environment at distance.	Good for safety, enabling a single user to control or interact with multiple physically distributed systems. Also, for use at distance for other reasons than safety, such as time to gain access, scale, extended observation or cost (such as extra-terrestrial, microscopic and marine exploration). Potential ethical issues of separation of operator from consequences of action and actions occurring under different legal regimes.
Teaming	The human and machine work together for a common goal (need be social).	Beneficial goals are facilitated (such as emotional components of user experience, efficiency and safety).
Symbiosis	The human and the system are closely linked working together for mutual benefit such as open, online courses and games in which the users get excitement and knowledge, and the system gets training.	Mutual benefit provided that both parties are aware of the agreements that they are entering into and do not abuse the relationship.
Parasitic	The human is a source of data collected by the system, but with little or no benefit to the human.	Not necessarily harmful but human not necessarily aware of RIA system action. This can be an issue that certain financial, political and organizational preconditions can foster, and that ergonomics should warn against. For example, designing an informed consent procedure.
Influence	Intelligent systems influencing human behaviour such as social media chatbots.	Behaviour can be influenced for the better. For example, safer. Regulation and licencing of pervasive, intelligent, automated propaganda and advertising.
Unknown alternatives	As yet undefined paradigms relating to organizational, social/cultural, societal relationships with RIA systems.	This is where the ergonomists need to cooperate with the experts on developmental psychology, sociology, business studies, law, etc.
Benevolent governance	Humans/humanity passing governance to AI.	For safety, collective benefit, reduction of hazards beyond human action. Agency is benevolent, having the best interests of society as the guiding principle. Agency provides governance: monitoring, evaluating, and directing plans and policies. Authority and safeguarding.

6.3.3 Perceived autonomy

From a human-centred point of view, apparently autonomous behaviour is a human-system issue and should be addressed in design. Indeed, from the human-system interaction perspective the degree to which system behaviour is perceived as autonomous can be a more important measure than the technical degree of autonomy of a system.

Since 2015, the terms "autonomous" and "autonomous system" have been widely used in marketing and journalism as synonyms for "automatic", "automation", and to describe systems with perceived autonomous behaviour. This usage is so widespread that it is not possible to assume any more precise meaning even in technical literature or regulation. As a result, behaviour that is perceived by a user as "autonomous" is a more common phenomenon than an autonomous agent.

6.3.4 Control loop

The control loop consists of the fundamental control functions, supporting information, affecters and effectors necessary for making adjustments in a process. A user, autonomous agent or other component of a system is said to be in the loop if that user, autonomous agent or component is in the process flow for making adjustments to the process. With the advent of RIA system agents, finer descriptions of where in the loop the user is located are emerging. For example, users who are in the loop during the execution of an automatic or autonomous process have interactions of a predominantly supervisory nature. An example of a less prevalent change in the use of in the loop is users who are before the loop. Such users have interactions prior to execution of an automatic or autonomous process but have no responsibilities for monitoring or making adjustments during the execution of the process. For a discussion of the decision/action cycle (OODA loop), see 10.3.

7 Categories of human-RIA system issues

7.1 General

The identified human-RIA system issues (problems) fall into six categories. These are summarized in Figure 1 and described in 7.2 to 7.7. Examples of ISO standards that can be related to these categories can be found in Clause 9.

This clause presents data for analysis, not conclusions. It describes human-system issues that are specific or particularly relevant to RIA systems rather than those that are common to non-RIA systems.

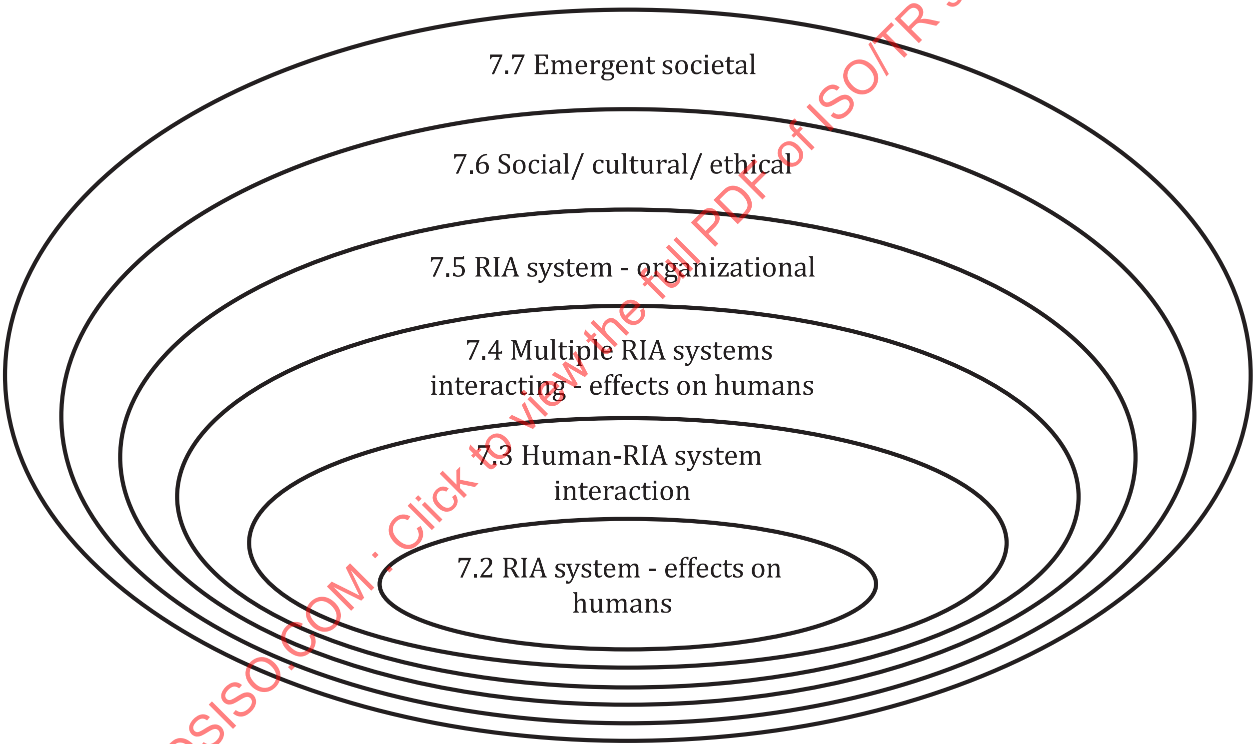

Figure 1 — The six categories of RIA system issues identified

7.2 RIA system — effects on a human

This first category represents the impact that the characteristics of an RIA system have on individual humans who are in the environment where it is operating. This includes physical, cognitive, affective, behavioural and motivational responses.

It should be stressed that the issues in this category are not limited to those experienced by users (those interacting with the RIA system or for whom its effects are intended). Unless an RIA system is totally unobtrusive, it impacts non-users as well, who have no vested interest in what it is doing. With respect to the non-users, an obtrusive RIA system in their environment can alter the way they perceive their environment, prompt an emotional response and alter their behaviour. For example, the mere presence of an RIA system can be perceived as a distraction or as an invasion of privacy or personal space. Therefore, one cannot expect the same effects to occur for users and non-users.

See A.2 for more examples of issues associated with how RIA system designed characteristics affect humans.

7.3 Human-RIA system interaction

Unless an RIA system is completely autonomous, humans interact with it directly at some level or to some degree. The humans in this category constitute users of the RIA system. This category is where most ergonomics resides today. It addresses the nature of tasks and the design of work to ensure that users are able to accomplish their intended tasks. The issues in the category specifically cover the consequences and impact of the design of the user interface on user (i.e. individual or team) interactions with the RIA system.

See A.3 for more examples of issues associated with human-RIA system interaction.

7.4 Multiple RIA systems interacting — effects on humans

One can anticipate that the time will come when multiple RIA systems operate within the same environment, coordinating activity and interacting without human intervention. Interactions can be obvious to the user, in the case of a physically embodied RIA system, or hidden, as can be the case with software agents, where the user only perceives a single system rather than its constituent elements. Where RIA systems from different suppliers or organizations interact, compatibility and communication issues between one RIA system and another are foreseeable, as well as with the humans present in environments where multiple RIA systems are operating. It is difficult to predict what emergent behaviours will result and their effect on users.

See A.4 for more examples of issues associated with the effects on humans of interacting RIA systems.

7.5 RIA system — organizational

The RIA system is likely to affect the activities such as work that occur in the organization, organizational processes, and the roles that people in the organization play. These issues raise the question of how we can optimize the existing organizational structures/working practices, etc., to make best use of a new RIA system in an existing organization. Although the issues in this category may not arise for every RIA system, they are likely to arise for any RIA system that is implemented at the level of an organization.

See A.5 for more examples of issues associated with RIA systems at the organizational level.

7.6 Social/cultural/ethical

RIA systems will not operate in isolation, but within a social and cultural context. An effective RIA system will seamlessly integrate within that context, as a familiar member of the team or a cooperating entity. Behaviours of an RIA system should be expected, fitting social and cultural norms. The issues in this category describe how the design of an RIA system can impact the social interactions, group attitudes and motivation, and collective group behaviours positively or negatively, depending on how seamlessly the system fits within the social and cultural context.

See A.6 for more examples of social, cultural, and ethical issues.

7.7 Emergent societal

The widespread or strategic use of RIA systems will re-define humanity's relationship to technology. In the longer-term, intelligent agents can be employed to make important decisions about governance. Assurances about the data and algorithms used, and how they are governed is likely to be a concern.

There will be sociotechnical issues related to how responsibility is allocated when large-scale intelligent agents are introduced. It is possible that humans will only be assigned work that compensates for incomplete autonomy, possibly resulting in roles that are not enriching, engaging, etc., and in which humans are not maintained sufficiently in the loop to be effective to step in if needed. There can be changes in employment brought about by the replacement of knowledge-based jobs with automated systems, in a similar manner to industrialization replacing skilled manual workers in the 19th and 20th centuries, but with the additional consequence that the higher-level cognitive skills that are lost can take generations to reproduce, should the need arise.

See A.7 for more examples of emergent societal issues.

8 Ergonomics and RIA systems

8.1 General

This clause introduces ergonomics as a way of tackling human-RIA system issues. RIA systems will be more connected, complex, probabilistic and non-deterministic, social, and augment human capabilities well beyond merely replacing physical work. Interaction with these agents can become a relationship, their interface a personality, and users and agents can form complex human-machine teams, working together towards a shared goal.

The benefits and problems identified in 8.2 and 8.3 will exist in the context of humans interacting with non-human systems. Ergonomics continues to be relevant, but with changing proportions of methods and techniques appropriate to these new RIA systems (see Clause 7).

RIA systems require consideration of the entire system design. The changes in the nature of tasks and the design of work to accommodate the complex, social human-machine interaction of RIA systems are fundamental for ergonomics. These changes will require that ergonomics addresses the systems approach, adapts its best practices and expands into areas of psychology and other social sciences that ergonomists do not deal with on a regular basis at present.

8.2 Benefits of ergonomics applied to RIA systems

Ergonomics has improved the usability of mechanical and computer controlled automatic systems. There are benefits to be had if an ergonomic approach is taken in support of the design and introduction of RIA systems. Systems and products with high levels of human-centred quality tend to be more successful both technically and commercially. RIA systems developed and introduced following human-centred principles will offer several more benefits. These include:

— improved ability for user/human to take control when required;

— increased accessibility for people with a range of capabilities;

— improved user experience;

— improved acceptance and integration of RIA systems in human teams;

— improved integration and support of human tasks and activities; and

— improved system performance across all RIA system states [primary/normal, degraded, emergency, and reversionary (when the system is unavailable)] and during the transitions between these states where particular issues can arise.

That can lead to:

— increased acceptance of the product/system/service;

— increased confidence and trust in the product/system/service;

— reduced supporting costs;

— a competitive advantage; and

— reduced risk of human errors and human machine teaming errors, unintended system use by users, and systems performing sub-optimally or entering unsafe or undesirable modes of operation.

8.3 Hazards if ergonomics is not applied to RIA systems

RIA systems present hazards for the individual and for society. The potential is high for negative impacts on the performance of individuals using the system and the behaviour of those otherwise affected by the system. These negative impacts include:

- inappropriate levels of trust/confidence in the system (having the correct level of trust or confidence in an RIA system is important. This is task-, context- and situation-dependent, i.e. an RIA system can be more reliable in some situations than in others, and the ability of a user to understand this is critical, through transparency/predictability of system behaviour, etc.):
  - under-trust can lead to: increases in workload due to additional checking and supervisory activities, users not adopting the system (if they have a choice) or being reluctant to operate it, and anticipated benefits of the RIA system not being fully realized;
  - over-trust can lead to errors being made or missed, task disengagement, and dissatisfaction with the system when it does not perform as expected;
- loss of situational awareness;
- disengagement from a task;
- over-reliance on the RIA system to complete tasks;
- increases in mental workload;
- generalized stress due to uncertainty created by the presence of an RIA system;
- degradation of knowledge, skills and abilities which can be required in reversionary, emergency and degraded system states or in the absence of the system;
- increases in training needs;
- perceived loss of control over one's work or one's environment, which can impact on job satisfaction and staff retention;
- ambiguity or loss of accountability and responsibility; and
- poor usability.

These negative impacts on the performance of individuals, in turn, can compromise safety by:

- increasing use error, leading to accidents and injuries;
- increasing the time taken for users to re-engage with a task (re-entering the loop) when they are required to do so, potentially delaying corrective action in time to prevent an unsafe state or action from occurring; and
- causing users to deliberately (or unknowingly) develop ways of working around the RIA system and violating processes in order to complete tasks satisfactorily.

Similarly, when the human element is not sufficiently considered, there are negative social impacts for groups of people using an RIA system or functioning within the context of an RIA system. These include:

- general dissatisfaction with the social aspects of the environment;
- disruptions in the functioning of teams using the system, or aberrations in social interactions due to the presence or functioning of the RIA system;
- general diffusion of responsibility in the case of negative events related to the system;

— perceived invasion of privacy by social groups and/or user teams;

— perceived decrease in security by social groups and/or user teams;

— generalized stress due to uncertainty created by the presence of an RIA system; and

— degradation of knowledge, skills and abilities which would take considerable time for a society to recover should the need ever arise such as electromagnetic or geomagnetic pulse.

These negative social impacts, in turn, can result in rejection of the RIA system, negatively impacting both the system manufacturer's and user organization's reputation, product sales or service performance and ultimately its bottom line or efficiency. They can also serve to create a dysfunctional workplace, with all the difficulties and costs that this entails.

Therefore, it is essential that a human-centred approach (human enhancement) be considered at all levels (of a system, organization and scale of system), where RIA system and overall system design is focused around the decisions and tasks that retained by the human user (across all potential use cases: optimal, degraded, reversionary and emergency) with the RIA system supporting these. The human should be kept in the loop to the extent that is needed to achieve required levels of performance and safety across all system states. This is particularly important when systems transition between states such as from normal operation into an emergency state. Determining the needed extent is part of system risk assessment.

Similarly, it is critical to seamlessly integrate the RIA system with the social context in which it operates, to ensure that its operation does not disrupt the social and cultural fabric of the environment. Thus, it is important that effective change management procedures are implemented whenever an RIA system is introduced into an environment that includes teams or other groups of people whose social interactions can be affected. Although change management is important during the introduction of any new or modified system, it is especially important for these systems to help ensure their acceptance and minimize disruptions of group activities in the environment.

9 Areas of RIA systems addressed by ergonomics standards

9.1 General

Standards for ergonomics provide information that is useful in addressing many of the issues described in Clause 7 and hazards described in Clause 8, particularly those related to user interface, organizational issues and lifecycle processes. Although these standards may not adequately cover the emerging/new requirements of RIA system characteristics, this clause summarizes how these standards can be used to address RIA systems.

9.2 Principles of ergonomics

ISO 26800 describes the general ergonomics approach and specifies basic ergonomics principles and concepts applicable to the design and evaluation of tasks, jobs, products, tools, equipment, systems, organizations, services, facilities and environments. The underlying principles of ergonomics remain the same, although the relative emphasis placed on them will vary. The principles and concepts described in ISO 26800 are fundamental to the design process wherever human involvement is expected, in order to ensure the optimum integration of human requirements and characteristics into a design.

There are standards on ergonomics and human factors based on these principles and concepts which can be used by executives, managers, engineers and designers in selecting, designing and managing systems and equipment to ensure that they are effective, efficient and satisfying to use. These are outlined below.

ISO 27500 draws on that extensive body of ergonomics and human factors knowledge and presents the rationale and general principles of human-centeredness in a concise form for executive board members and policy makers. It explains the principles which characterize a human-centred organization. These principles are as follows:

— capitalize on individual differences as an organizational strength;

— make usability and accessibility strategic business objectives;

— adopt a total system approach;

— ensure health, safety, and well-being are business priorities;

— value employees and create a meaningful work environment;

— be open and trustworthy; and

— act in socially responsible ways.

They provide a framework for organizational behaviour when RIA systems are considered or implemented.

ISO 27501 outlines the responsibilities of managers in supporting a human-centred organization, in fulfilling each of the seven principles with reference to internal, external, and societal stakeholders.

9.3 Human-centred design process

Human-centred design is an approach to interactive systems development. By applying human factors/ergonomics, and usability knowledge and techniques, human-centred design aims to make systems usable and useful by focusing on the users' needs and requirements. This approach enhances effectiveness and efficiency, improves human well-being, user satisfaction, accessibility and sustainability; and counteracts possible adverse effects of use on human health, safety and performance.

ISO 9241-210 provides requirements and recommendations for human-centred design principles and activities throughout the life cycle of computer-based interactive systems. It is intended to be used by those managing design processes. It is concerned with ways in which both hardware and software components of interactive systems can enhance human–system interaction. The approach described complements existing systems design approaches. It can be incorporated in approaches as diverse as object-oriented design, and waterfall, agile, and other rapid application development processes.

The principles and activities of a human-centred approach to design are elaborated in two human-centred design process models. ISO 9241-220 describes the processes that ensure human-centred quality of interactive systems. ISO/TS 18152 describes the processes that address human-system issues in the engineering of systems. The processes in these standards go a lot further than the current norm in HCD practice (i.e. specifying processes to support HCD in governance and project management, saying that HCD people get involved early, defining the user requirements and then driving the system/technical/platform level requirements) and will apply to RIA systems as much as to other types of system and any set of human-systems issues. ISO 9241-220:2019, Annex F, provides guidance on risk management and human-centred design.

9.4 Interaction and interface

ISO 9241-110, ISO 9241-112, ISO 9241-13, ISO 9241-129 and ISO 9241-154 set general design principles for the ergonomic design of interactive systems, and the information is presented without reference to situations of use, application, environment or technology. They provide a framework for applying those principles to the analysis, design and evaluation of interactive systems. While they are applicable to all types of interactive systems, they do not cover the specifics of every context of use such as safety critical systems and collaborative work. The principles are generally independent of any specific design style or application. ISO 9241-129 specifically addresses adaptive dialogues. Other parts address novel and multiple modalities.

A dialogue is the "*interaction between a user and an interactive system as a sequence of user actions (inputs) and system responses (outputs) in order to achieve a goal*", where user actions include not only entry of data but also navigational and other (control) actions of the user. Principles relating to dialogue and individualization help prevent usability problems including:

— additional, unnecessary steps not required as part of the task;

— misleading information;

— insufficient and poor information about the user interface;

— unexpected response of the interactive system;

— navigational limitations during use; and

— inefficient error recovery.

Principles related to information and guidance provide a variety of benefits to users including improvements in speed, accuracy, mental effort and user experience. They also help prevent users from experiencing usability problems with presented information. Examples of such problems include the following:

— not detecting information despite it being present;

— other information distracting from information the user is focusing on;

— not discriminating between pieces of information, since they appear to be identical;

— misinterpreting information, since the meaning of the information is ambiguous;

— expending unnecessary time in understanding information, since the information presentation is unnecessarily lengthy;

— unfamiliar content and format leading to lack of understanding of information.

ISO 9241-154 gives guidance on, and requirements for, the user interface design of interactive voice response (IVR) applications. It covers both IVR systems that employ touchtone input and those using automated speech recognition (ASR) as the input mechanism. It is equally applicable to cases in which the caller or the IVR system itself such as in some telemarketing applications initiates the call. It is intended to be used together with ISO/IEC 13714, which covers voice messaging specifically. Although ISO 9241-154 covers many of the same topics as do the other standards cited in this clause, ISO 9241-154 is application-specific, and provides guidance that extends beyond general principles.

9.5 Accessibility

The ergonomics definition of accessibility adopted by ISO focuses on broadening the range of users of products, systems, services, environments and facilities by taking into account the widest range of user needs, characteristics and capabilities to enable as many users as possible to achieve their goals in particular contexts of use.

ISO 9241-171, ISO/IEC 29136, ISO 9241-129, ISO 9241-971[1)], and ISO/TR 22411 provide guidance on specific aspects of the design of an interactive system that take account of the widest possible range of users. The application of the specific guidance takes place within the application of the human-centred approach to design, as described in 9.3. EN 17161 describes an approach to design taking account of accessibility, which focuses on the inclusion of the widest possible range of users.

ISO 9241-171 provides ergonomics guidance and specifications for the design of accessible software for use at work, in the home, in education and in public places. It covers issues associated with designing accessible software for people with the widest range of physical, sensory and cognitive abilities, including those who are temporarily disabled, and the elderly. Most of the provisions in ISO 9241-171 apply to any part of the software that implements or contributes to the software user interface.

1) Under development. (Stage at the time of publication: ISO/DIS 9241-971.)

ISO 9241-171 points out that in addition to adopting a human-centred approach to design, the capacity to provide individualization is an important tool in increasing the accessibility of a human-system interface. This enables diverse user needs, characteristics and capabilities to be met by adaptation and tailoring to specific requirements. This can include automated identification and adjustment of interface and interaction characteristics in response to specific user behaviours. ISO 9241-129 dealing with individualization is therefore discussed here, as well as in 9.4. In addition to describing specific implementation features, ISO 9241-129 provides general recommendations on when and how it is appropriate to adopt individualization.

ISO/IEC 29136 provides requirements and recommendations for the accessibility of personal computer hardware, to be used when planning, developing, designing and distributing these computers.

ISO 9241-971 provides ergonomic guidance focused on ensuring the accessibility of tactile/haptic interactions with ICT systems. This includes both general and specific guidance for the design of accessible tactile/haptic interactions. This document deals with modality shifting and accessibility guidance for haptic/tactile modalities such as gestures, vibration and force feedback.

ISO/TR 22411 does not provide guidance on design, but provides data sets which describe the functional characteristics, capabilities and responses of diverse users. In particular, it has extensive amounts of data relating to older users. The data can be used to determine applicable values and dimensions which are appropriate when designing for the widest possible range of users taking account of variations in the specific contexts of use.

9.6 Workspace and workload

ISO 6385, ISO 9241-2 and the ISO 10075 series provide ergonomic principles for the design of tasks, work and work systems. They encourage attention to human, social and technical requirements in a balanced manner during the design process.

The systems approach in these standards assists in both existing and new situations, such as introduction of an RIA system. Ergonomic evaluations of existing or new work systems will show the need for, and encourage attention to, workload and the role of the worker with/within those systems.

Work system covers a large variety of working and leisure situations, including permanent and flexible work places. Work systems involve combinations of people and equipment, within a given space and environment, and the interactions between these components within an organization. Work systems vary in complexity and characteristics.

The principles specified in these standards support design of optimal working conditions with regard to task performance, workload, human well-being, safety and health. This includes the development of existing skills and the acquisition of new ones, while taking account of technological and economic effectiveness and efficiency.

Technological, economic, organizational and human factors affect task performance, behaviour and well-being of people as part of a work system. Applying ergonomic knowledge in the light of practical experience in the design of a work system is intended to satisfy human requirements consideration of workload and its measurement.

9.7 Context and environment

ISO 9241-11 provides a framework for understanding the concept of usability and applying it to situations where people experience or use interactive systems (including RIA systems), and other types of systems (including built environments), and products (including industrial and consumer products) and services (including technical and personal services). Usability is a scalable, task-based measure of the degree to which users are enabled to achieve goals effectively, efficiently and with satisfaction, taking account of the context of use. ISO 9241-11 explains how usability can be interpreted in terms of human performance and satisfaction. It emphasizes that usability is dependent on the context of use (the specific circumstances in which a system, product or service is experienced or used).

Control and control centres present an early and widespread application of RIA systems. The overall strategy for dealing with user requirements in control centres is presented in ISO 11064-1. ISO 11064-2 provides guidance on the design and planning of the control room in relation to its supporting areas. Requirements for the layout of the control room are covered by ISO 11064-3. Ergonomic requirements, recommendations and guidelines for the design of workplaces in control centres are established in ISO 11064-4. Displays and controls, human computer interaction and the physical working environment are presented in ISO 11064-5 and ISO 11064-6. Evaluation principles are dealt with in ISO 11064-7.

ISO sets standards for modelling, measuring and assessing the impact of properties of the physical and thermal environment. As RIA systems increase their presence in environments, especially when controlling dynamic integrated environments, personal capability, comfort and safety will increasingly depend on machine application of these standards and correct interpretation of the integrated effect of environmental factors.

10 Changes in ergonomics standards required to better address RIA system technology

10.1 General

The human–computer interaction (HCI) paradigm (inherited from ergonomics) is largely task-based. Tasks are defined by action steps or goal achievement. So far, HCI has focussed on task support such as affordances of an artefact. This is because focus on action leads to focus on supporting that action with tools. Considering RIA systems, we need to shift focus from supporting users of tools to facilitating teams achieving goals (teams include single or multiple users and single or multiple intelligent agents). Furthermore, rather than an emphasis on completing a task, the focus is shifting to engaging the user and shaping behaviour, on shaping the nature of work. Human-centred quality (HCQ) still applies as a measure.

Risk and goal-based approaches to standardization are likely to work better than prescription. This fits with the approach to specification for the ISO 9241 series. But experience tells us that they are harder to explain and implement and their use carries a level of uncertainty with which many stakeholders are uncomfortable.

10.2 Type of guidance needed and for which readerships

Ergonomics traditionally involves the design and evaluation of the user interface and its context. The ergonomist develops and studies the concept of operations for a system, defines the tasks to be performed, and allocates those tasks to the user, the machine, or some combination of user and machine. The physical presentation and layout of controls, the content and format of information displayed to the user, and the physical workspace are of interest to the ergonomist to ensure accessibility and usability of user interfaces. Sensation and perception, cognition, learning and memory, social and organizational interaction and teaming, human performance assessment, ergonomics and biomechanics are general examples of knowledge, skills, and abilities common to ergonomists.

The toolbox of principles and techniques applied by the ergonomist to develop user interfaces in a system has evolved to support systems that function according to inputs provided by the user in context with engineered models of the environment in which the system is to be operating. Automation runs on such models, often following scripted logic for performing tasks the user would have otherwise had to perform.

Ergonomists need to continue to be the advocate of the user by applying a broader range of tools/approaches, from, for example, sociology, anthropology, social and organizational psychology, emotional design and user experience (UX). The toolbox of principles and techniques applied by the ergonomist will need to be modified to address the unique interaction challenges autonomy presents. As automation extends to become more complex autonomous, dynamic, probabilistic and non-deterministic behaviours that are more adaptive to the operational needs of a system will be necessary. Merely focussing on the design and layout of control panels will be insufficient to define the user interaction with the system. Human-centred design approaches are needed, that:

— are consistent with, and linked to the wider systems engineering and design processes for RIA systems;

— are capable of dealing with the increased complexities of systems-of-systems – and the implications that they have for users;

— account for appropriate calibration of trust in the RIA system, and assist the users in deciding when and to what degree is necessary to over-ride the RIA system and subsequently taking on such tasks;

— facilitate HF requirements identification, architecting, modelling, design, human in the loop tests, acceptance processes, etc., for RIA systems.

Integration with engineering, management and design, and user requirements management, are already addressed in ISO/TS 18152 and ISO 9241-220.

With advanced autonomy that functions like another person or team, collaborative awareness and crew resource management are examples of constructs and techniques from teaming and supervisory command and control that can apply. Likewise, the management of abnormal situations can also provide relevant guidance.

Like the current use of stereotypes and affordances, understanding biases in social expectations and social contracts of a specific user population will be essential for the developers of autonomous systems to consider and control for in their designs. This can be as simple as ensuring the proper use of turn signals in an auto-driving car, or as complex as deciding the level of medical care to be given to an elder.

Ergonomics standardization consideration of issues, principles and techniques relevant to ergonomists dealing with the development of autonomy needs to address that middle ground between what we know about how to design effective interfaces between people and automatic systems, and the industrial and social psychology of system implementation, teaming, and supervisory command and control. Other domains, like sociology and anthropology, can have principles and techniques that can prove useful to ergonomists dealing with autonomy; as can computer science.

10.3 Transparent interaction and transparent users

For humans and RIA systems to work together via transparent interactions in the future, not only will the user need to understand a system's current state, but also the functions that it performs in the activity being conducted, and also predict the behaviour of the system under given circumstances/conditions. Likewise, the system will need to "understand" the users' current state, the functions that it performs in the activity being conducted and predict the behaviour of the user under given circumstances/conditions. This implies that both user state monitoring sensors and models of the task, user goals and human behaviour have been built into the system. This is likely to be complex due to the wide variation in human characteristics, behaviours, cultures, experience, physiological and cognitive state, personality, etc. Without accurate (or at least appropriately accurate) models of humans and systems, it is possible that a human will have to adapt to, accommodate and compensate for sub-optimal system behaviour/decisions resulting from inadequacies in the machine models. If the RIA system is designed to learn (i.e. to modify its models of humans and systems), care should be taken regarding how what was learned is (or is not) to be incorporated into the system and applied. Humans are not always free to change how they perform a task when they discover what can be a more efficient or otherwise superior process for performing the task, particularly when functioning on teams or otherwise interoperating with others, e.g. Figure 2.

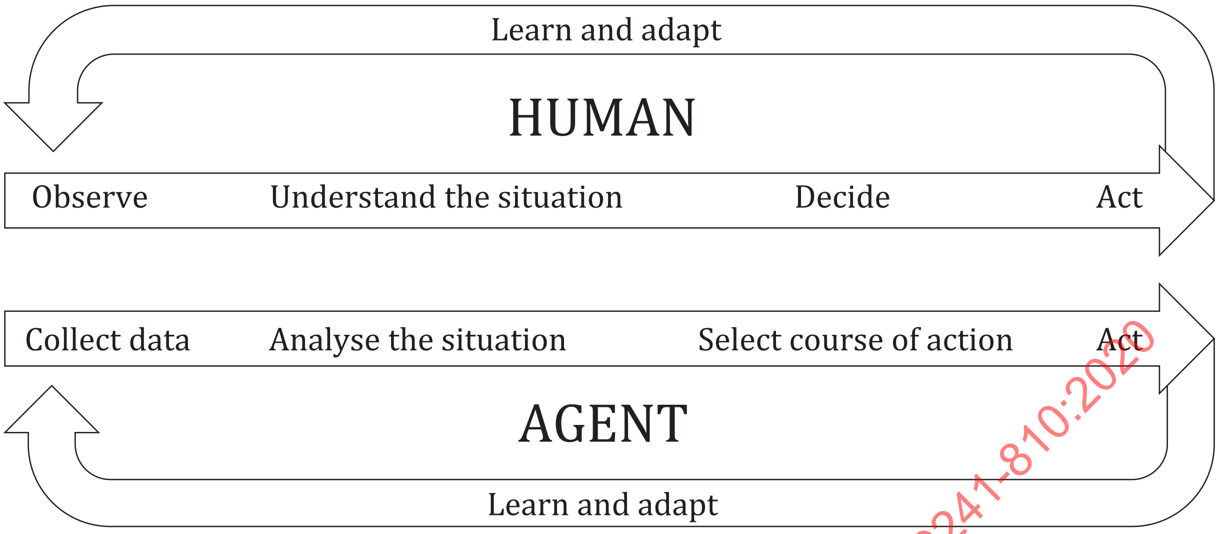

Figure 2 — Human-agent teaming dual Observe–Orient–Decide–Act (OODA) loops

To optimize human-machine teams, it will be important that both the human and the machine elements are working towards a common goal with a shared understanding of the current situation and desired end state. Without a shared understanding of both the goal and the current situation, it is possible that the human-machine team will become inefficient or even have human and machine working in conflict with each other. It can be necessary to develop complimentary task models to help facilitate the development of these transparent interactions between human and machine in RIA systems.

10.4 Safety aspects of RIA systems

RIA systems can introduce new safety issues. For example, something that temporally is too fast for humans to address in real-time, or something that is otherwise not directly observable such as online activity. Latency, the time that humans have to take over from an RIA system that detects that it has lost control, is another factor. Or, at a societal level, there is the "Frankenstein issue": the systemic consequences when the inventor (human or humanity as a whole) does not take timely responsibility for a created entity. Safeguards are necessary, but what form should they take and what part does ergonomics play?

A safety set is a fundamental set of normal/expected human behaviours in context such as robots at edge states of physical engagement with human bodies. These would probably be phrased as sets of safety requirements ("*the system shall do this … under these circumstances*"), or cases. For example, the safe condition for a self-driving car to achieve in the event of loss of control is: parked on the hard shoulder with hazard lights on. For online human-systems safety, it can be, for example: "do what a user does most frequently"; but issues, such as offline actions not accessible to machine learning, can lead to different, dangerous interpretations by an intelligent agent such as donning personal protective equipment. In other cases, what is considered ethical can be most appropriate. Table 2 presents a human-centred analysis of safety issues against the human-system issue categories.

The safety set relates to ethics/laws/regulations. It is not easy to decide what is in a universal set. The user/RIA system relationship approach (as presented in Table 1) is likely to have an effect on both safe action and construction of safety/harm. Different safety sets can be required for different cultures. A cultural construction of safe/safety should be added to the context of use. There is probably a different safety set for every category, although there will be interrelated causes and effects between categories (for example, attitude to robots extends to social). To facilitate the consideration of safety, ergonomics standards for RIA system should include a means of assessment. Consideration should be given to the broadest application of measurement of human-centred quality as a basis of assessment. This includes avoidance of harm, which is a human-centred treatment of safety and other types of harm (both to and from humans). ISO 9241-220, ISO 25063 and ISO 25065 contain supporting material on HCQ, context and user requirements respectively. ISO 9241-220:2019, E.5, contains examples of harm from use.

Table 2 — Safety aspects by category of human-system issues

Issue category	Type of safety set	Location of requirement	HCQ emphasis/ development
RIA system – effects on humans	Effects on individual humans that are to be prevented	Regulation (part of rights and occupational health and safety)	User experience
Human-RIA system interaction	Situations to be prevented, acceptable latency, authority to decide/override	Performance standard, possibly mandated (similar to competence)	Usability, accessibility, avoidance of harm
RIA systems interacting – effects on humans	Degree of transparency, ability to override	Regulation, business practices, engineering standards, contract	Avoidance of harm
RIA system – organizational	Process requirements, governance, effect on stakeholders, learning behaviours	Industry sector codes, and standards of practice, possibly mandated	All components as objectives plus context coverage
Social/cultural/ethical	Fairness, trust, offence, safeguarding, privacy	Regulation or convention, possibly faith ruling, probably tort	Extension of user experience to address community experience Extension of harm to address cultural issues
Emergent societal	Loss of control, accountability	International conventions and regulations	Application of HCQ at societal level

Annex A
(informative)

Human-RIA system issues

A.1 General

The identified human-RIA system issues (problems) fall into six categories. These are summarized in Clause 7. They are described in more detail in this annex. It presents data for analysis, not conclusions. It describes what human-system issues are different for RIA systems, not what is the same. It presents the simplest explanation of what we can come up with at present that calls out the new, specific issues under the category.

A.2 RIA system — effects on a human

See 7.2.

Social, cultural, ethical and societal issues are addressed by the categories described in A.6 and A.7.

Introduction of RIA systems is often undertaken with the intent to aid human activity and improve task performance, thereby improving reliability, efficiency, and throughput for the affected processes. It can readily support individualization and account for and utilize the emotional impact of such interaction during task performance. It can assist human activity on tasks requiring high physical and mental task load, memory and attention. It can facilitate accessibility to services that existing manual task requirements make difficult. RIA systems can free human resources from having to physically perform dull, dirty, or dangerous tasks, such tasks then becoming significantly less sensitive to fatigue and injury. Likewise, tasks that require a long time to perform and consistent precision over the course of that time can also benefit from the adroit application of an RIA system, making it less sensitive to fatigue and shift changes of the remaining human operators. The shifting of tasks to autonomy and the increasing of monitoring and management tasks of the remaining human operators does not necessarily decrease task loading and workload for these operators. Likewise, awareness of a) what the system expects of the users (feedback from the system to the user) and b) what the system is doing and why (status information) can be adversely impacted.

Machines are presently considered to be tools. Their design characteristics afford use by human operators to facilitate or enable the precise, repeatable performance of a specific, often highly constrained task or set of tasks. Though machines can include automatic systems that precisely and repetitively perform the same set of tasks according to mechanical design or software scripts, they have typically not been considered to have characteristics of highly adaptive partners during the performance of tasks. RIA systems have designed characteristics that have only been possessed by other human operators or team members. These characteristics include significantly greater freedom of action by the machine and the strong potential to adapt and learn how to better perform the tasks assigned to the machine, loosening constraints on how the machine that includes an intelligent agent needs to act to accomplish these tasks. An RIA system can act without active user input regarding how the system is to function, focusing instead on what task(s) the human operator expressed a need to be accomplished. An RIA system can employ personality as an attribute of the interface, leveraging social interaction to elicit desired inputs and effects to accomplish what the system has been tasked to accomplish. These characteristics are known to have an impact on the ability of human operators or team members to effectively perform desired tasks of the team. The more distal, collaborative or supervisory nature of task allocations to the human operators of an RIA system can result in a degradation in knowledge, skills and abilities for the tasks performed by the system. Situational awareness of the details of the performance of the task and/or an elevation of cognitive workload can diminish, shifting instead to a collaborative awareness. As RIA systems take on more complex tasks that would otherwise require multiple human actors, advancing what any single human actor can do, the ability of an RIA system to nominally function with significantly fewer human operators can result in an inability of the human operators to resume such tasks in the event of system failure. If personality is not an embedded feature of the design of an RIA system, anthropomorphism can result in the human users misattributing personality to the system, leading to inappropriate or otherwise misleading assumptions about how the system will react to attempts at interaction.

A.3 Human-RIA system interaction

See 7.3.

This set of issues addresses how RIA characteristics change the form of interaction between human operators and an RIA system. The ramifications of imbuing machines with intelligence and independence of action present a significant shift in how human operators interact with machines to accomplish tasks. As it was learned when automatic control was introduced to machines, such reallocation of tasks from human to machine not only resulted in significant changes in how the human operators need to be supported in the system, but also resulted in many new and emergent human tasks that did not previously exist.

RIA systems can permit new modes of interaction with non-human autonomous agents, e.g. adaptive dialogues, higher-level combined modes. Human behaviour itself can form part of the interaction with an RIA system, therefore requiring consideration of behavioural norms for the population using the system. Implicit and explicit forms of communication, and affordances of the interaction space, are additional examples of interaction modes that RIA system will enable and require consideration of. social touch and other forms of physical proximity during the performance of tasks can be features of the interaction with an RIA system, opening up the user interaction space to include personal, social, political and existential levels.

The cultural and societal differences in social norms and ethics need to be considered by RIA systems, or violation of such norms can result in confusion regarding what is expected of the human user by the system, user discomfort regarding working with an RIA system, distrust of the system, and ultimately rejection of the technology. An RIA system that does not consider social norms in its interactions with the user can also result in incidents or accidents involving direct or indirect interaction with the system. Adaptation of RIA systems to the social norms of the user would require that an RIA system be able to decode social cues or personal characteristics and respond according to such norms. If RIA system behaviours are not designed in a human-centred way, feelings of irritation, of surveillance, a lack of controllability and transparency of the system, etc., can result. Trust in the system would degrade and become considerably more difficult to recover.

Given the potential for adaptive behaviour, bounding RIA system behaviours and revealing those bounds will become important. Conversely, the use of an RIA system to enforce safe behaviour, i.e. removing an element of choice and behaviours of the human user, can serve to reinforce security, health and safety. This capability of an RIA system to act against the control inputs of the human user potentially results in ambiguous command authority of the user at best. How do we communicate to users what they are in control over and what the system is in control over, especially when this can change dynamically or unique accessibility needs are presented? What methods should users have to change and keep track of levels of human/RIA system control? Interpreting the intent of an RIA system can be challenging. Without strong, timely, implicit/explicit communication, affordances and appropriate/expected behaviours by the RIA system, users may not be able to accurately interpret the intent and expectations of the system. The ability to interpret intent couples with trust, in that a difficulty or inability to interpret intent can reduce trust in the system. Likewise, an RIA system can need be able to perceive and understand how the behaviour of the user impacts performance of the task and the goals that the user is working toward. If an RIA system does not correctly interpret the intent of the user, the RIA system and user can work at crossed purposes. With greater interconnectivity, this misinterpretation of what the user is trying to achieve can spread throughout a system, leading to inefficiencies, poor performance, frustration, higher workload, etc. These aspects of human-machine teams raise the potential need for:

— humans to have appropriate mental models of the current and likely future RIA system state and behaviour;

— the RIA system to be sufficiently transparent to allow users to understand current system state and behaviour;

— the RIA system to have appropriate models of human behaviour and capabilities;

— the RIA system to have a means of monitoring human state and behaviour;

— a means of human and system synchronizing or aligning their situational awareness, understanding and task goals.

A.4 Multiple RIA systems interacting — effects on humans

See 7.4.

This encompasses system-of-system issues but can identify additional user requirements to the engineering perspective. Alternatively, users can perceive responses from separate RIA systems as connected/related. This perceived interconnectedness can have negative consequences. For example, if one element of the system is unreliable a loss of trust in one element of functionality can spread to other, reliable elements of the system affecting adoption/acceptance of the system as a whole. Typical system-of-system cases are:

— alarms from several devices in an ICU, wherein the nurses and doctors would need all measurements from one system to solve the problem;

— ship bridge alert management system, including human monitoring/awareness of RIA system to RIA system interaction;

— mode interactions and transitions between an aircraft flight management system and the autoflight systems;

— conflicting behaviours between the RIA systems under the same circumstances, e.g. the interaction between the blood oxygen meter and dosimeter used by an anaesthesiologist.

RIA system interactions can occur within the context of manufacturer-defined protocols that are not exchanging information with the human users. When RIA systems are allowed to interact and evolve their communication, human users may no longer be able to understand the interaction and unanticipated emergent behaviours can result. Likewise, when RIA system interactions occur between systems designed for two very different cultures, misinterpretation or other confusion can result when RIA systems use their human interaction protocols for the communication. RIA system to RIA system interactions can also require that the human users coordinate the interaction or act as a bridge between incompatible systems.

When multiple interacting RIA systems are supporting or otherwise assisting a user in the performance of a task, the user is still responsible for the performance of the system-of-systems. Collaboration awareness includes:

— an understanding of the presence and features/capabilities of the combined system that convey member activity and task-relevant status;

— an understanding of what and where things are happening with the various agents and their shared resources and products;

— an understanding of how things are progressing relative to the goals driving the use of the RIA system.

A.5 RIA system — organizational

See 7.5.

Organizations can be broken down into several types:

— supplier organizations that design, develop and support the RIA system life cycle;

— organizations that acquire and leverage the abilities of RIA systems to achieve a commercial, societal or other goal;

— organizations or population groups that are affected by the fielding of an RIA system, e.g. a workforce, customers, general public, patients, etc.

There is likewise a distinction between issues related to implementation organization vs a design side organization vs user organization. Implementation organizations own the requirements for the particular needs defined by the user organizations. Design organizations are in charge of the characterization of RIA systems and support the value chain. Design organizations can be defined in terms of developers of the components (e.g. the robot, ML package), and an integrator that ties together components to create a system. User organizations provide support and maintenance for an RIA system, and training for users. User expression of requirements are largely through user behaviours, purchasing decisions, explicit expression, etc., that are engaged with RIA systems. Infrastructure that exists (e.g. regulatory organization) defines what can and cannot exist in society, licensing or certification of RIA systems, and regulatory and legal issues for user requirements. This includes organizations at the higher societal and infrastructure levels, where enterprises combining many organizations need to connect in some way to realize and sustain new ways of living.

Issues in this category predominantly address the reinforcement of safety, risk compensation, domain/operations knowledge and responsibility/liability. Inserting an RIA system into a workplace presents similar challenges as automation has, requiring adaptations to be made that risk a suboptimal workplace for the human workers. Responsibility and accountability issues need to be understood early in system design. They typically impact organizations, for instance, by creating new organizational functions for monitoring and troubleshooting ("cleaning up after the bot"). Organizational structures and business/production/working process and practice will evolve to make best use of a new RIA system in an existing organization, as will identification of opportunities for continuous improvement and the use of the data generated by the RIA system working within the organization. See also 9.3 regarding lifecycle HCD processes.

A.6 Social/cultural/ethical

See 7.6.

An RIA system will consider the social and ethical norms of its actions within the society in which it operates (e.g. behavioural expectation, tolerance, fairness, compassion), and how violation of such norms can result in confusion regarding what is expected of the human user by the system, discomfort regarding the use of the autonomy, and ultimately rejection of the autonomy. Autonomy that does not consider social norms can result in incidents or accidents involving human direct or indirect interaction with the system.

RIA system technologies are being applied to military, industrial and civilian applications. Crossover of technologies can be beneficial (e.g. paramedic and medical systems, security systems), but they can also be detrimental (e.g. the use of civilian RIA system technologies for malicious purposes in criminal or terrorist applications). The required level of training for users and the norms for behaviour can shift significantly when an autonomous and/or robotic system technology is applied to another domain.

Introduction of non-human autonomous systems will:

— alter existing task allocations and result in new and emergent human tasks;

— require reconsideration of authority, accountability and responsibility for actions taken;

— change the roles of humans and their associated teams and organizational structures in the workplace, with wider societal impacts.

In warfare, remotely piloted systems are permitting shirt-sleeve combat and days that begin and end from the comfort of one's home and family life that are divorced from the violence of combat. In civil society there is a risk that unless wider human issues are considered during the introduction of an RIA system the automation or reallocation of tasks that provide job satisfaction can result in workplace dissatisfaction, unrest and recruitment and retention issues.

The deskilling of jobs can also result in a reduction in the resilience of the organization or process to RIA system failure, or unanticipated external events.

RIA systems separate the human from the performance of tasks, increasing the moral distance from the consequences of actions, making it more likely that tasks that can result in negative consequences will be performed as more capable autonomy is introduced, e.g. someone declining a loan to a customer, or refusing health insurance because of the recommendation of an RIA system, removing human discretion and moral latitude. The willingness of a person to take risks regarding actions with moral consequences increases with:

— greater perception of distance from the action;

— more indirect nature of the tasks leading to the action;

— greater anonymity.

This phenomenon would also exist for those who train the RIA system, raising the question of whose persona will be reflected in the future behaviours of RIA system. If an RIA system is allowed to learn and grow on its own during the course of its interactions with users, those experiences will likewise shape the future behaviours of the system.

Benevolence of users cannot be assumed. A decrease in empathy and an increase in anti-social behaviours with respect to the society in which the RIA system is functioning can result if an immutable check on fundamental behaviours is not established and enforced on the RIA system.

A.7 Emergent societal

See 7.7.

Issues tend to emerge during or soon after a new scientific or technological development is applied. These issues generally involve society as a whole, and are dealt with via social change, policies, standards, regulations and procedures. Thus, although human factors/ergonomics specialists can have a role to play in addressing some of these issues for RIA systems, many other types of professionals, as well as the general public, will work together in addressing these issues.

The boundaries of ergonomics, HCD processes and methodologies, applicable theory, scope of influence, design principles and measurement methods and metrics can need to change as the current boundaries of ergonomics are expanded to accommodate the social aspects of RIA systems. Likewise, existing standards for other aspects of computer systems can begin to overlap with standards for ergonomics, particularly regarding taxonomy, testing, assurance, specifications, and system integration. Context of use can prove particularly important for RIA systems given the social aspect of such systems. Testing, assurance and specifications, in particular, can present challenges with RIA systems given the non-deterministic/probabilistic nature of such systems.

Systems that change behaviours over time (e.g. machine learning) can present a particular challenge for the design of the human-machine interface with an RIA system. Without strong implicit/explicit communication, affordances and appropriate/expected behaviours by the RIA system, users may not be able to accurately interpret the intent of the system. The ability to interpret intent couples with trust, in that a difficulty or inability to interpret intent can reduce trust in the system. Equally, without appropriate feedback from sensors and an understanding of the task and goal that the user is working towards, the RIA system may not correctly interpret the intent of the user, resulting in them working at crossed purposes. With greater interconnectivity this misinterpretation of what the user is trying to achieve can spread throughout a system leading to inefficiencies/poor performance/frustration/higher workload, etc.

Annex B
(informative)

Examples/case studies of ergonomics issues for RIA systems

B.1 General

A variety of examples and case studies of ergonomics issues for RIA systems are provided here. These include autonomous systems in general, with brief examples from aviation and other automobile transportation systems; industrial robotics; medivac; online support (especially using IVR) and activity management and conflicts between human and RIA working environments. Additional examples not discussed in detailed here but identified include: information seeking/libraries, communicating at distance, maintaining something, travel, surgery, sharing news, entertainment, caregiving, assistive technologies (e.g. cochlear implant, prosthetics, exoskeleton, spell/grammar check, electronic maps) and balancing workload by temporarily engaging/disengaging automation.

B.2 Autonomous systems

Users understand their current workspaces because such spaces are designed with affordances and other implicit and explicit communications presented by the interface that are common to the users' experience and flow naturally with the task and the environment in which the task is being performed. Ergonomists design such workspaces with the intent that the spaces be consistent, intuitive and predictable for the target user population. Ergonomics principles and techniques have been developed with this and more in mind. When a user approaches such a workspace, the user knows what is expected of him/her to accomplish the task supported by the workspace.

What happens when the workspace can autonomously adapt, when the workspace can change its behaviour to more efficiently accommodate the needs of the user to accomplish supported tasks? What happens when a previously autonomous workspace degrades, offering less assistance, less support for the user? The problem shifts from developing and reinforcing the user's awareness of a consistent workspace, i.e. one in which the user knows exactly what is expected at any given moment, to developing and maintaining the user's awareness of what is expected of them now in a dynamic context.

Mode confusion is one example of an adverse effect that can happen when a system automatically both changes its states and the logic for changing those states. Supporting the construction and maintenance of the user's awareness of what the automation is doing takes up considerable time in the design process and subsequent user training. Even with good design, the user sometimes has to expend considerable effort in constructing and maintaining awareness of what it is that the autonomy will do and what is required of the user.

The user's awareness of the context of a system change and the ramifications for the user's prediction of future behaviour can present an additional challenge when systems change their behaviour in response to stimuli. Existing automatic systems have a rule-based algorithm in their back end that can cause system changes but can still be comprehensible for the user. Whereas systems which embody something like a strong AI, an algorithm that learns its own rules, can cause system changes that are not comprehensible at all (e.g. Go-Play, where the system came up with the best solution but nobody knew why). This second type of control in an autonomous system can lead to even stronger effects of mode confusion and non-transparency.

The significant perceptual and cognitive burden of monitoring an autonomous system, and skill degradation of the user when the user has relied upon the autonomy without practicing the knowledge, skills and abilities required by the tasks performed by the autonomy, are additional examples of adverse effects involving the introduction of autonomy into our workspaces. Stress and fatigue in these examples magnifies the burden on the user given the greater demand on attention, decision making and memory that the operation of autonomous systems can require.

Likewise, relying on the user as a failsafe, dumping full operational responsibility onto the user when the autonomy fails, will present considerable challenges for control and information display for tasks involving complex, safety critical, potentially high tempo operations. There have been many examples in aviation, for example, where failure modes in the autoflight systems involving reversion to partial or full manual control of the aircraft have resulted in incidents and accidents. Less extreme and certainly more common in aviation, automatic mode transitions in autoflight systems that require differing responsibility for manual control from the pilots has also contributed to the incident and accident statistics. Considering the high degree of training and recurrent training of pilots, the incidence of mode confusions and inability of every pilot to assume full control over the behaviour of the aircraft in a timely manner is troubling. What will reversion and mode transition awareness do to the prevalence and severity of incidents and accidents in domains that don't require such stringent training requirements? The time dimension is essential in designing dynamic systems, where unpredictable or excessive time pressure is a major factor. Present day examples exist of operators engaging and disengaging automation to temporarily balance workload. A more general case is adapting use of automation. The emerging autodriving car technologies have already provided examples of autonomy failing with reversion to manual operation into the hands of a driver who was not ready to take full control over the car. With the more advanced capabilities autonomy represents, reversion can require compensation for a level of processing of essential information and assessment of options that a single unaided user cannot perform in the time available.

The guiding principles and solutions to these and many other challenges for the designer of autonomous systems are likely to extend the toolbox that an ergonomist typically employs to develop user interfaces in a system. Social psychology, organizational and industrial psychology, and anthropology are examples of fields of study that may offer tools and techniques that the ergonomist will find useful in the study and design of RIA systems intended for use by a diverse population across cultures.

B.3 Automobile applications

As suggested in A.2, systems should clearly show what mode they are in and whether the system is "live". A number of serious injuries have occurred in offshore drilling when automated systems activated because they can be powered with no visible sign and then move suddenly. The indication of power was in the control room but not on the equipment. Thinking about self-driving cars, should other road users know whether the car is in self-driving mode or under driver control? With respect to communicating intent, people are good at interpreting what other people are doing, particularly from their own cultures, but not as good for people from different cultures or for non-human entities assuming the tasks that people normally perform.

An RIA system should be able to be stopped by any person at risk. Or, where this is not practical or appropriate, it should clearly direct to a controller. For example, if a pilot encounters a UAV, they should be able to predict its behaviour as they would another manned aircraft. The ability to call the UAV pilot in real time if the UAV poses a risk is an appropriate consideration.

Another example from aviation is how pilots fly using a rail line to navigate, keeping the rail line on the left as is standard practice in some countries, so that there is not a head on collision with someone coming the other way. Drones are often programmed to fly directly over the line, which potentially reduces the margin of safety.

Regarding the issue of control and monitoring, how do we decide which systems require monitoring with the scope for rapid human intervention and which do not? In the case of a person in a self-driving car, should they be competent to drive or in a fit state so that they can drive out of harm's way in an emergency, or is this not required because they have no role in the system or legal duties? If monitoring is required, for example a central supervising agent oversees a number of units, then what are the requirements in terms of situational awareness and intervention times for the drive and for the people monitoring the system? Is this monitoring a reasonable task, or is it too difficult with a high level of concentration combined with long periods of nothing to do? In a 2018 fatal self-driving crash (Uber), the monitoring driver was not a young alert tech person but a rehabilitating ex-offender with a known substance abuse problem. The controls required by the driver to assume full control over the vehicle were disabled, negating whatever ability the monitoring driver would have had to assume positive control over the vehicle to avoid the deadly collision. This would suggest that monitoring jobs can be very hard to perform by some users if they can be performed at all, with a knock on to poor intervention times.

This Uber fatal crash has yet to be fully investigated at the time of this writing, but it raises another issue. The sensor manufacturers have already announced that their sensor would have picked up the person who was killed, so they are blaming the programmers who defined the hazard parameters for recognizing a person in front of the car. But the person who was hit was behaving in an unexpected manner by pushing a cycle laden with plastic bags in the road, resulting in an “unexpected” profile from the perspective of the programmers of the system.

It is interesting to speculate that the programmers had little experience of homeless people who push their belongings around on bicycles at all hours and so did not add suitable rules to the "safe to ignore" ruleset. The question raised here is, what life experience or diversity should be required in the decision programming to ensure that all the real-world hazards are covered? Is there a generalized control schema that can be applied that the automatic or autonomous system can fall back on to safely navigate around an unidentified or otherwise unusual threat (e.g. “aviate, navigate, communicate” from aviation)? The extent to which a system continues to rely to some extent on human operators, particularly when the system is presented with a situation that it does not know how to handle begs the question of how does the human supervisor or monitor know when something has failed or is about to fail?

RIA system that will operate in close proximity with people should have much greater understanding of context, particularly human behaviour around and with the product. How is the boundary defined between safe operation and unsafe operation? What is the time horizon for recognition of and response to behaviours? Auto-driving cars currently have large written exclusion lists, lots of "gotchas" because of context. Aviation presents what can be readily achieved today, but the operation of an aircraft with a high degree of autoflight capabilities requires a high level of investment and training necessary to safely accommodate such capabilities.

Perhaps a clearer explanation of what a device can do will be required. What is the "when all else fails" basic task? Does the designer know what the basic task really is? Does this extend to programming/parameterization? Can it be a basic set up? What are the lizard brain functions? An example here is for a lawn mower: "don't run over things on the grass".

If a system learns and develops expertise, how does one know what level it has reached? Regarding the Mercedes driver assistance systems, the transmission learns driver style. If a transmission is replaced, the new system is naïve and the driver is instructed to let the system re-learn. How would this work for more complex systems so that we would know the system is now expert and understands what this meant, for example that it can act autonomously in a situation that it used to require control input?

The highest level of authority held by the RIA system should be clear to all parties. There are emerging issues in the design of the decision algorithms for self-driving cars. Research is being performed regarding how to prioritize in a crash. For example, should the vehicle always protect the driver as they are the customer, even if that would mean driving into a group of pedestrians or cyclists? If the consensus is no, then a product is being sold that can decide to take an action that can kill its owner. Obviously, this is true of all vehicles today, but the fatal decision is made by the driver not the product. There is a fear that a black market will occur in which wealthy owners illegally re-map the collision algorithms or other protections to always favour themselves at all costs. Given that there is already an illegal re-mapping industry to circumvent emissions controls, this can be considered a real risk.

B.4 Industrial robotics

In the manufacturing setting, workspace design has previously not had much to consider regarding direct interactions between user and robot/autonomous system. Safety law and regulation demanded complete segregation of traditional high-speed, high-payload industrial robots via physical barriers and fencing. This strict enforcement was necessary for the very reason that it is not sufficient to rely solely on human monitoring and awareness to mitigate risk. Various safeguarding measures (sensor-based speed and separation, hand-guiding modes, power and force limits) now offer safe industrial human-robot collaboration/coexistence (as reflected in ISO 10218-1 and ISO/TS 15066). This is generating a significant increase in levels of small- and large-scale shop floor robotics and autonomous systems because integration and robotic process automation (RPA) can greatly enhance business performance. However, regardless of safeguarding measures, this closer interface impacts the design of working tasks themselves. It will inevitably bring greater human cognitive demands and need for awareness of autonomy mode as direct interaction between user and robot/autonomous systems becomes more common. Attributes of the working task such as the amount of decision latitude, skill discretion, timing, and method control a user has in a task can be directly affected by closer human-robot interaction.

Whereas the workforce previously did not have close proximity and direct interaction with segregated systems performing rudimentary tasks, operators on the shop floor are now working in collaboration with systems performing more complex tasks will have to interpret precisely what the system is doing, predict what it is going to do next, and decide what actions he or she may need to take. Moreover, these demands will be significantly exacerbated when the environment comprises multiple systems and operators with varying levels of autonomy and mobility, requiring collaborative awareness of state changes. Training and experience will help operators to develop new skills to perform their work tasks, but the design and integration of these systems should take account of cognitive ergonomics and psychological well-being and safety, not just technical/functional safety and business performance as it is currently.

B.5 First responder treatment of a casualty and transport on to hospital

This theoretical example considers the first responder treatment of an injured member of the public and their onward transportation to hospital. In this example, the user of the RIA system under consideration is the organization that provides the service. However, there are other potential users, such as the patient, personnel administering care or transporting the patient to hospital and the maintainers of the systems involved. Considering each of these different users can give alternative perspectives of the system and its impact, risks and opportunities.

In this example there are several phases involved in the response to an injured member of the public, each of which lend themselves to different types of RIA system support. These phases in this theoretical example are:

— tasking of the first responder to the injured person;

— initial diagnosis;

— initial stabilization and treatment of the casualty;

— transport of the patient to hospital;

— hand-over of the casualty to hospital personnel and onward treatment.

In this example, not all of these phases are considered in detail, but it is intended to highlight the range of potential human factors that can be considered.

In the past — On finding the casualty in the street, a member of the public would locate the nearest public phone and call a local emergency control centre, verbally providing information regarding the location of the casualty, before returning to the casualty. The control centre calls the ambulance station closest to the casualty and passes the information onto the ambulance station controller who passes the information to the ambulance crew in the station. The ambulance consists of a van containing a stretcher, crewed by a driver and a first aid technician. The ambulance crew navigate to the patient using local knowledge and paper maps if required. On arrival at the scene, basic first aid is administered, and the casualty transported to hospital, notes on the treatment provided and any additional information gained regarding the casualty is recorded on a paper record form. On arrival at hospital, a porter unloads the casualty from the ambulance and takes them into the emergency department with the paper patient record, which is passed to the hospital medical team.

In the recent past to present — On finding the casualty in the street, a member of the public uses their mobile phone to call an emergency number. The call is automatically routed to an area control centre based on the location of the caller. The location of the caller is provided to the emergency call handler who collects information from the caller to populate a pre-prepared, computer-based dialogue with information about the patient and their condition. The call handler identifies the nearest available ambulance to the casualty, which are distributed across the area in predetermined locations and verbally passes on the information regarding the casualty location and condition. The ambulance consists of a specialist vehicle containing a range of separate pieces of equipment for diagnosis and treatment and crewed by a first aid trained ambulance technician and more highly trained paramedic. The ambulance crew enter the casualty location into a satellite navigation system, which provides a recommended route to the casualty. While the ambulance is on route, the caller can stay in contact with the call handler who provides advice, based on a first aid training course, on basic treatment that can be provided to the casualty. On arrival at the scene, the paramedic uses the diagnostic equipment to make a decision on what treatment to provide. The paramedic decides at what point the casualty is in a stable enough condition to be moved and transported to hospital, at which point they are loaded into the ambulance and transported to hospital while monitoring and treatment of the casualty continues to be provided by the paramedic. Patient information and details of any drugs and treatment administered are recorded on a paper record. On arrival at hospital, the patient is unloaded by the ambulance crew and a verbal handover to the hospital team is provided together with the paper patient record.

At present and into the near future — The same basic concept outlined above is maintained, but with the addition of support technologies that affect certain areas of the response.

— The information collected by the call handler and entered into the computer system is used to populate a decision aid that provides suggested step by step first aid interventions that the call handler can pass onto the person at the scene to conduct on the patient and/or request additional information about the injuries/patient. This can reduce the training requirements of the call handler and/or provide a wider range of potential treatment options as well as improve the quality and consistency of the instructions given to the member of the public at the scene.

— Ambulance locations and status are continually tracked in real time. When the need for an ambulance to attend a patient is identified by the call handler, the system determines the best ambulance to dispatch based on proximity, priority and near real-time traffic information. The system passes a tasking instruction to the ambulance together with the electronic information collected by the call handler.

— Ambulances can be dynamically re-tasked on route should higher priority cases arise or if the system identifies a more optimal tasking solution across all of the active calls and traffic information.

— Rather than pre-determined, fixed ambulance locations, ambulance positions change dynamically based on historical emergency call data and other real-time factors such as traffic density, weather conditions, events occurring, etc.

— Ambulance carried diagnostic and treatment systems are interconnected, exchanging patient data in order to provide the paramedic with a central source of patient information and decision support in order to help identify and select appropriate treatment options.

— Patient data is transmitted to the emergency department in advance of the patient arriving, an RIA system seeks additional information about the patient (if they are known) from a central medical records database and provides synthesized/fused information together with potential diagnosis and treatment options to the medical team that will be caring for the casualty.

In the future — The same basic concept outlined above is maintained, but with the addition of more advanced support technologies that affect certain areas of the response.

— Patient data is collected at the scene by passers-by using smart devices (potentially the patient's own worn device) and passed to an RIA system and human team coordinating the initial response. Patient data is combined with medical records and fed into the RIA system which makes the diagnosis and treatment plan which is then implemented by the paramedics attending the patient.

— Augmented reality systems allow remote specialists or RIA system to view the patient and provide in vision instructions to either a member of the public, guiding them through simple first aid steps, or to the attending paramedic guiding them through more complex procedures.

Further into the future:

— The patients' smart device alerts the hospital that the user is in need of critical and immediate care. Real-time health data is passed to the RIA system and human team responding to the patient.

— An unmanned ground vehicle (UGV) transports paramedics to the patient and then to hospital allowing both paramedics to administer care on the move. With additional casualty handling systems and changes to some equipment, only a single paramedic can be required.

— The paramedic team attending the patient can also be made up of a human and a medical robot, which can be autonomous or remotely/tele-operated by a specialist in the hospital to conduct more complex treatments.

— If required, a UAV can transport a doctor rapidly from a hospital to the patient in order to deliver specific care.

In the far future:

— A UAV carrying an un-trained human porter flies to the casualty a swarm of medical robots administers treatment with the human porter supporting the robots to do the things that they cannot, e.g. repositioning the patient. The UAV transports the patient to the hospital, administering care while on route and the porter is transported back via a UGV taxi deployed to collect them.

— The patient's smart devices identify the early signs of a condition requiring treatment and advises the patient to take a UGV to the nearest emergency department. Their smart device provides continuous monitoring and dispatches an emergency response UAV/UGV to rendez-vous with the patients UGV for emergency treatment and transport to an emergency department.

— Tele-operated or RIA treatment systems distributed across an area allow comprehensive emergency treatment to be administered quicker and without the need to a centralized emergency department. Doctors travel by UAV to the patient in the treatment units for onward treatment and care.

B.6 Intelligent agents and assistants for customer service

Early in their evolution, one of the driving forces for intelligent agents was to reduce the need for human customer service representatives and to encourage self-service. One of the first implementations of intelligent agents was in interactive voice response (IVR) systems, with user interactions via touch-tones or automatic speech recognition (ASR). Users were, and continue to be, dissatisfied with many of these applications for a variety of reasons, some of which relate to the state of the technology itself, but some of which relate to the design of the user interface.

Through one means or another, these applications require the users both to make queries and respond to system queries in order to execute a transaction or receive desired information. Normally, when a human being asks a question, they do not think about the exact words or the grammar they use. They expect to speak naturally, and this continues to pose a problem for many systems for which the natural language understanding capabilities are limited. Through experience, the human user may learn how to formulate acceptable queries (i.e. simple ones that the system can handle). This increases the mental work involved in using the system and decreases the perceived human-ness of the system, which many users feel is important. They want to have a natural language conversation with the system, and when the system cannot accommodate them, they will often abandon it. Such systems violate the prime directive of human factors which is to fit the task to the human, not the human to the task.

User difficulties often occur because many of these applications use little of the sources of contextual information they have in order to interpret user queries, and they also make little use of external sources of information that are available, such as a calendar or e-mail information, to make reasonable inferences about the user and his/her desired objectives. Sometimes, they make assumptions that frustrate user goals. Many of GPS-based route selection and navigation assumes the user wants to minimize either travel time or distance. If the user's goal, on the other hand, is to travel a scenic route, the GPS will actually thwart any attempt to do so be re-routing the user at every turn, to meet the goal the system has assumed.

Although standards exist for the design of user interfaces for IVR systems, including for dialogue design (e.g. ISO 9241-154), these standards are seemingly not employed in the design of many RIA systems. This leads to numerous user interface problems, including users having to input the same information more than once in the course of an interaction or within a series of interactions, because the databases and other information sources being used by the system are not sufficiently integrated. Since the 1990s, most IVR applications have been and continue to be examples of RIA systems. IVR systems used for customer service represent some of the first RIA systems to become ubiquitous. Although the impetus for their development was specifically to remove the human customer service provider from the communication loop, few IVR systems, if any, meet this goal in practice, due to poor user interface design, and the challenge of managing all contingencies even for relatively simple applications. The human-human interface thus continues to exist in these applications and requires human factors attention, to the same degree as does the human-technology interface.

Finally, methods for evaluating the success or failure of these systems, from the perspective of the user, are often not good, because there can be a huge difference between the goals of the user and the goals of the company or organization that has implemented the application, and insufficient effort is made to determine how often user goals are achieved when using the system. As an example, in many IVR systems, the company's goals are to minimize call time, and decrease the number of calls transferred to a live agent. Some companies do not realize that neither of these measures necessarily reflects user success or failure in achieving goals. A hang-up or application closure can indicate user success, but it can also reflect a discouraged, dissatisfied user who has simply abandoned their attempt to obtain service. With more and more of these systems providing no option, or only an exceedingly difficult-to-use option, for reaching a human agent, it is ever more important for companies to have good performance metrics for assessing user satisfaction and customer goal attainment.

A recent article summarizes the current state of intelligent assistants as follows: "Usability testing finds that both voice-only and screen-based intelligent assistants work well only for very limited, simple queries that have fairly simple, short answers. Users have difficulty with anything else." While television advertisements and demos make these systems look impressive, the reality is a bit different.

B.7 Social networks/online activity management

As the concept of the personal computer via the maturing of network technologies transformed into social networking, developers and design researchers started to articulate their experiences from successful applications into guidelines in the form of design patterns. This semi-structured formalism for presenting best practice in context declared problem-solution pairs allowed for practical design recommendations that would support values such as transparency, mutual trust and community building. The naming of a pattern should explain its usefulness with expressive and story-like wording about how a module or component should be designed in a user-centred way. This can be exemplified from an early compilation of patterns for welcoming a new user into a social networking platform: quick registration, login, welcome area, mentor, virtual me, user gallery, buddy list.

Now, with the introduction of machine learning (ML) packages in the commercially very successful social networking area, designers and developers need to be aware of how other values can influence user experience. People are not the data about them. Identity is not completely defined by the observable behaviours of a person. At best, a straw-person persona can be developed by observations that can be made online. Intention is not adequately captured by ML looking at search parameters, shopping habits or a stated political position on social media. Commercial and political value considerations can be embedded in different modules and components, thereby creating tensions with the user-centred values that originally were the driving forces when the developers themselves were among the influential users. User guidance patterns (from the same collection), like Birds of a feather, which compares user profiles and suggest collaboration with other users, can now also have to take account of parameters and functions in algorithms supporting user profiling, comparison and collaboration that can be irrelevant, or even contradicting to what users would expect. The same concern need to be paid to rewarding patterns, such as Hall of fame, which lists users who have helped other users the most to contribute to community goals.

Among community building features in social network platforms that can be vulnerable to the cumulative effects of ML influences are what has been called shared browsing, application sharing and embedded chat, three design patterns aimed at supporting friendly and trustful sharing of information, application and messaging support. Again, when such ML effects become oblique, either because of the sheer complexity of the ML algorithms or by commercial or political intent, a renewed attention needs to be paid to user guidance and consent.

The themes of user guidance and consent are foregrounded in a recent contribution on patterns for designing RIA applications. Without going into the nuts-and-bolts of how transparency can be supported algorithmically, the pattern names evoke a sense of the challenges ahead for the designers, developers, platform administrators, managers and owners of social network applications and platforms: show the man behind the curtain, open up the black box, and demonstrate fair and equal treatment.

The patterns exemplified above are from one of the challenges identified in the RIA pattern collection, assuring users perceive good intentions. Other challenges identified are protecting privacy (e.g. data security is the foundation, tailor expectations to context), establishing successful and long-term adoption (e.g. always ask: who is being made the hero, plan for the role of human resources), and demonstrating accuracy and reliability (e.g. explain the conditions of accuracy, prove success by showing failure).

Given the growing commercial and political interests invested in RIA applications, and the exponential complexities arising when this breed of systems becomes networked, it is difficult to envision how the often idealistic concerns and arguments of design researchers and media critics can find their way into a public discourse about the challenges ahead. However, the arguments articulated in the design pattern collections quoted from here stem from concerned designers and developers within the software industry. And the recent critique regarding privacy breaches by some of the dominant social platforms points towards a growing awareness from regulators and politicians, as that soon will materialize in demand for new standards. Because the pattern collections mentioned, and others to be collected and publicly scrutinized, are often concerned with the same set of values as those promoted, for example, in ISO 26000 and ISO 27500, they provide an important input for future standardization efforts.

B.8 Potential conflicts in working or living environments in which accommodate both humans and RIA systems.

Where humans and RIA systems operate or co-exist in the same physical or information environment, there may be conflicts between the needs or desires the of humans working or living in that space and the requirements of the RIA system to operate effectively and efficiently. For example, a warehouse optimized for automated pick and pack robots may not be designed to accommodate the anthropometric characteristics of the maintainers required to conduct periodic maintenance, resulting in a compromised working environment for humans in the system.

Within the constraints of an industrial workplace good design, training and working practices can help avoid or manage potential conflicts between humans and RIA systems developing a solution in which both can operate safely, effectively and efficiently. However, within the public environment this can be more complex where different people and RIA systems operating within or moving through the environment will have different needs, objectives, behaviours and expectations of the world around them. Some of these public requirements can be functional and common to both people and systems (for example, being able to move from one place to another safely and quickly) while others can be non-functional and purely human (for example, aesthetic, cultural, social, etc.). Within urban design and town planning, there will be a need to consider how to optimize public spaces for both RIA systems and humans to minimize the potential for conflict or detrimental social impacts while maintaining the benefits that RIA systems will enable.

Does one:

— optimize the environment for RIA systems and rely on humans to adapt to it? It may not be cost-effective to design RIA systems to deal with the complexities of the human environment;

— design the environment for humans and design the RIA system to operate safely within it? This runs the risk of very complex expensive systems that do not deliver anticipated benefits and risk failure at the numerous edge cases;

— design an environment that is not optimized to either and accept lower levels of performance from the RIA and the human?

This conflict is not constrained purely to the physical space. In the information domain, the objectives of an intelligent system can be at odds with those of the user. For example, an online shop's customer support chatbot which a customer is using to answer a query regarding the future availability of an out-of-stock item can have two objectives: the first to answer the customer's query without the need for an expensive human call centre, the second to encourage the user to spend money on the site now, by using information regarding the customer's previous buying habits and personalized incentives. With increasing information availability and personalization of these systems, apparently helpful interactions can lead to electronic coercion or manipulation of the human user by the system.

While this is not new, the sophistication of intelligent systems can make them more discrete or effective. In this use, case ergonomists will need to consider two users of the system, the first being the retailer who wishes to make a profit from consumers and the second the customer using the site. Ergonomists can be asked to apply knowledge regarding human behaviour and decision making to systems design to influence, manipulate, coerce or at worst deceive end users. This dual user paradigm has existed within the retail sector long before the advent of the internet, through good customer service delivered by shop staff. However, as intelligent systems become more prevalent the potential of moral, ethical and legal aspects of system design cand become more complex.

Annex C
(informative)

Development of ergonomics

C.1 The view from the top — emergent issues for traditional ergonomics in developing RIA system

There is active research relating to the human interaction with RIA systems spanning almost the entire breadth of the human sciences. However, in focussing on those aspects of research and emerging trends that are of most relevance to the scope of this document the following are some emergent areas which future standards in this area can need to consider in more depth:

— human machine teams (HMTs) — In the future, RIA systems will act more like teammates than tools. Therefore, the design of future systems will need to consider not just the design of the system to support effective and efficient task-work, but also the fostering of effective team work between human(s) and machine(s). The design of future systems will have a significant impact on the speed with which these HMTs mature and perform. Therefore, the human factors community will need to:

  — understand the factors that impact on HMT maturity and how best to measure it;

  — understand how systems design impacts HMT performance, speed of maturation and what the characteristics of good HMTs interface design are;

  — provide guidance on how to design systems that support both task work and team work;

  — provide guidance on supporting motivation, joy of labour and the other emotional aspects required for HMT performance;

  — assess whether there are new risks and error types that emerge from HMTs;

— systems-of-systems integration and emergent properties — While there are many challenges associated with optimizing discrete human-RIA system interactions, there are potentially significantly more complex human factors issues that arise when multiple RIA systems and humans, forming HMTs, operate in the same physical or information environment. There are likely to be complex socio-technical issues and unpredictable emergent behaviours, functions and risks arising where these systems meet and interact. These will need to be carefully managed during design and introduction into service of new systems and new human factors analysis and modelling tools to support this can be necessary. This is likely to be particularly challenging where:

  — systems are designed to adapt their behaviour in order to optimize their own performance (as in the case of ML) or the strategic goals of the owning organization. There is the potential for conflict where several adaptive systems and humans, with differing goals and purposes interact, leading to undesirable outcomes in performance and safety terms;

  — systems with different levels of autonomy have to interact;

  — RIA systems designed for different cultures, organizational structures and legislative rule have to interact.

In order to address this, there are likely to be new human roles within systems to manage these emergent properties. This will present particular human factors challenges in terms of:

— how humans (at individual, team organizational and governmental levels) maintain awareness of the behaviours and states of the RIA system and exert control over them;

— identifying potential forms of system (human and RIA system) error;

— temporality — RIA systems will be able to operate and interact (physically and cognitively) at a speed that humans are unable to match. This raises particular challenges in terms of how humans interact with systems containing multiple interacting RIA systems. Challenges are likely to include maintaining awareness of system state, the ability of users or organizations to regain control (or change the level of human control of the system), identifying and mitigating error and unsafe states;

— system resilience — Historically, when automation fails or is unable to handle a particular situation a human operator is expected to step in to regain control. Within the design of RIA systems that are intended to support or remove human cognitive functions, there is a particular risk that the human user will not be sufficiently situationally aware, engaged in the task or have the cognitive resources available to be able to retake full control of the system in sufficient time to avoid a negative outcome of unsafe state. To address this, it will be necessary that HF practitioners are able to support the design of systems that deliver satisfactory performance across all potential system states (normal, degraded, reversionary, emergency) and operating scenarios while also striving for optimal task performance during normal operating conditions;

— effective human control — For some tasks/situations, there will be a need to ensure that a human remains in effective control of the system. This can be required for legal, moral or ethical reasons. Understanding how to design and assure systems that require effective human control is complex and will require a multi-disciplinary approach;

— adaptive systems — Some future RIA systems will dynamically adapt to the cognitive and physical state of the user, reallocating functions between the human and the technological components of the system and/or changing the level of support provided in order to maintain whole system performance. In order to deliver this type of capability, systems will need to include both task models and models of human capabilities and responses to task load, environmental stressors and other influencing factors. The human sciences community will need to develop these models, as well as robust means of systems collecting the data required to populate them from the users of the systems.

C.2 Later steps, further development

In summary, the areas in which ergonomics will need to develop to fully address RIA systems in the future are listed below:

— stakeholder analysis for RIA systems (e.g. team members, who is the user, who are the stakeholders, multiple RIA systems interacting with the team and each other);

— social aspects of RIA systems and effects on society (e.g. highly social interaction, cultural ergonomics, personality dimensions for RIA systems, organizational psychology/ergonomics, management of emergent behaviour);

— human state measurement and models (e.g. human state measurement, human models, general models of humans, models of specific users/user groups);

— human machine teaming (e.g. dynamic function allocation between humans, machines; dynamic goal setting; dynamic task allocation between human team members; dynamic task allocation between machines; transparent interaction and transparent users; knowledge acquisition/management; small data vs big data – learning/evolving systems vs models);

— guidance and RIA systems (e.g. transparent systems, automated education, training users of RIA systems, design of systems for zero training);

— principles for RIA system ergonomics.

C.3 Safety principles

To refine the contribution of ergonomics in this difficult and potentially dangerous area will require understanding and collating known, new and emergent types of human error and human machine team error in the context of the different types of RIA system, e.g. human errors resulting from loss of situational awareness during periods of high automation and low workload.

Areas for study include the following:

— unwillingness to challenge flawed RIA system decision making/action, including update of human error analysis techniques (e.g. HEART) and team resource management (e.g. CRM) to include RIA system issues;

— speed of action — The speed with which an automated or autonomous system can react to a potentially hazardous situation has many benefits (e.g. collision avoidance systems). However, this also introduces new safety issues. For example, where an RIA system is acting at a speed which is beyond the capability of a human to monitor or intervene, there will exist the risk that human intervention to prevent the system from performing an unsafe or inappropriate act (either in the physical or information domain) will not be possible. This is more likely in edge case conditions or unanticipated situations which were not considered within the design of the RIA system. It is possible that as systems move towards their design limits their speed of action may need to be throttled to allow human intervention if required;

— observability of action — Where an RIA system is operating in the information domain, it may not be possible for a human to visualize/conceptualize the complex range of interactions and the first, second and third order effects that actions have in this domain.

Annex D
(informative)

Changes required to ergonomics standards

D.1 General

The standards for ergonomics described in Clause 9 already address many of the issues described in Clause 7 and hazards described in Clause 8, particularly those related to user interface, organizational issues and lifecycle processes. However, this is not sufficient to ensure the correct and consistent consideration of human-system issues for RIA systems. This subclause discusses where standards for human-system interaction and related standards will need to be developed to better address RIA systems. The standards to be developed are discussed under the categories described in Clause 7.

It is not a surprise that there are more and more general topics in the higher number issue categories. System effects on individuals and at least single user interaction are more likely to have been addressed in standards.

Each clause quotes the category description and lists the aspects of standardization that need to be developed to address the human-system issues that emerge for the category. Subclauses outline possible developments. They elaborate on the areas for development by category described in Table D.1. Table D.1 outlines the ergonomics aspects for each category of human-system issues and suggests the standards that can be developed to address these aspects and to be used in the testing/evaluation/licencing/certification of the human-centred quality of RIA systems. These are speculatively numbered as an ISO 9241-81x series, on the basis that it is easier for an established team to take on this work rather than attempt to pass it over to another group. It is hoped that this work stimulates the consideration of further standards.

Table D.1 — Ergonomics aspects by category of human-system issues

Issue category	Aspects that require consideration	Suggested standards and Gaps
4.2 RIA system — effects on a human	— Environment (physical effects) — Personality (expectation for systems that do more than tools, relationships, anthropomorphism), — Workload (adding distraction, cognitive, skill degradation) — Context (with effects on person) — ISO 9241-100, environment — what and how to test? — Human state measurement and models Architecture of guidance is a set of constructs/rules for dealing with situations which need to be detailed in context, but we should define a foundational set for each of the following based on current principles of system design: — Acceptable (safe) outward behaviour — Acceptable proximity, appearance, and language — Interaction behaviours — Tribal (group) identity — Attitude towards risk — Humour — Interrogation — Attention — Level of knowledge of AI	ISO 9241-811 Ergonomics design considerations for RIA systems Things to consider in the design of an RIA system user interface with parts on: — software: natural language and contextual aspects of communication (e.g. character, culture, medium, safety, trust, emotion(al), ethical, appearance); — physical: non-verbal communication, robots, presence, augmentation (e.g. gesture, appearance, safety, expectation); — environments: virtual and physical, control (e.g. aesthetic, comfort, preference, safety, wellbeing, expectation, immersive). ISO environmental ergonomic standards cover the physical and comfort aspects. Standards relevant to robotics industrial safety currently focus on the potential effects on the physical safety of a human that may be caused by a robot/machine system and its constituent parts. Clauses address functional design and implementation measures but effects on (or from) human behaviour and psychological responses are not addressed (or considered relevant).

Table D.1 *(continued)*

Issue category	Aspects that require consideration	Suggested standards and Gaps
4.3 Human-RIA system interaction	— Latency — Work/task design/new task — New modes of interaction, adaptive, physical (UI space list) — Software ergonomics, — How testing/ evaluation is changed — Human machine teaming — Guidance and RIA systems — Training design	ISO 9241-812 Ergonomics of interaction with RIA system (parts for current paradigms): — replacement (levels of automation); — teaming (e.g. crew resource management, authority, levels of trust); — augmentation (e.g. allocation of function, feedback). This would probably take the form of principles, methods, and guidance, for each level or by type of risk. For robots, as above. Interactions of human teams in the RIA system are not addressed.
4.4 Multiple RIA systems interacting — effect on humans	— Explanation of action — Environment (combined effects) — Understanding — Safeguards — Crew resource management (mixed teams)/mediation. RIA system as users. Testing or design?	Need for research into, issues related to: — safety; — operations; — extension of system of systems. Also need to clarify the relationship between ergonomics and systems engineering. For robots, as above. Potential effects from RIA systems (multiple RIA agents) on humans are not addressed.
4.5 RIA system — organizational	— Organizational change to realize benefits — Design of work — Stated and observed requirements — Regulation assessment and certification — Context of use (e.g. more social) — Changes to organizational structure — Use of information — Change management — Stakeholder analysis for RIA systems	ISO 9241-814 Process issues for RIA system — Necessary extensions to ISO 9241-220 — Potential revisions of ISO 27500 and ISO 27501 Organizational and management issues For robots, as above. Organizational impacts and influences are not addressed.

Table D.1 *(continued)*

Issue category	Aspects that require consideration	Suggested standards and Gaps
4.6 Social/ cultural/ethical	— Context of use (of social and cultural norms) — New or radically changed human tasks (including value of these tasks) — Distance (e.g. physical, moral) — Norms of behaviour — Effects on resilience — Trust — Evaluation (what? how? paradigms?) — Extension of HCD process? — Social aspects of RIA systems and their effects on society	ISO 9241-813 Context of use for RIA systems (e.g. more on social issues in previous column) Potential modifications or extensions of ISO 9241-11 and ISO 26000 For robots, as above. Social/cultural/ ethical impacts and influences are not addressed.
4.7 Emergent societal	— Overlap with other standards (internet of things, artificial intelligence, integration, testing of use) — Analyse and explanation of human-system aspects — Sociotechnical methods, new standards with sociotechnical aspects (e.g. employment, communication of learning/capability, decision transparency) — Definition of accountability and responsibilities	Need for research to support regulation For robots, as above. Emergent societal impacts and influences are not addressed.

D.2 RIA system — effects on a human

D.2.1 General

See 7.2.

Aspect that require consideration are: environment (physical effects), personality (expectation for systems that do more than tools, relationship, anthropomorphic), workload (adding cognitive, skill degradation), context (with effects on person). (ISO 9241-100, environment), what and how to test?

D.2.2 Workload

Need to add integration work, emotional workload, new forms of stress, new patterns of workload, effect on required competence, particularly in context of reduced manning, development and maintenance of situational awareness, changes in duration of attention/required reaction time.

D.2.3 Human state measurement

The way in which systems monitor user state (cognitive, physiological, behavioural) and task performance and use this to inform their human models of intent, fatigue, workload, task engagement, emotional state, attitudes etc underpins a wide range of the ergonomics areas in this section. Dynamic function allocation will require this, and many of the more social aspects of RIA systems will rely on them.