Deep Science Publishing
https://doi.org/10.70593/978-81-989050-0-0_11

Chapter 11: Infrastructure for autonomy: Designing networked service systems with intelligent interfaces

11.1. Introduction

Human autonomy will increasingly depend upon well-designed networks of artifacts and infrastructures. As the world becomes ever more complex and uncertain, our ability to shape it in virtuous ways will rely upon the extent to which diverse resources — people, natural resources, biological systems, software, and hardware — can be organized to serve our needs directly and concurrently. The most promising technological developments affecting current solutions for these foundational requirements of human existence are in networked information systems (Gubbi et al., 2013; Chen et al., 2018; Dautov et al., 2020). The emergence, development, and eventual proliferation of the Internet and its offspring will increasingly enable solutions for the urgent challenges facing modern societies. However, this incredible diversity of intelligent, networked devices, operated by sophisticated communication protocols, is complicated to comprehend and use. Modern societies demand solutions that transcend the traditional solutions of organizational theory; by optimizing a mix of human beings, information machines, and artificial interfaces, it is now feasible to offer real-time support and continuous augmentation of human decision-making and deliberation. Information technologies can finally minimize the burden of coordination among people who have a relevant problem to solve, or a service to provide, and who have to interact to achieve a common goal. Blueprinting models of autonomy for societies cannot yet account for the massive changes in individuals' decision-making patterns induced by novel technological solutions for communications, transport, and assistance. New processes of commoditization — and new modes of orchestration — of human talents are emerging and will profoundly reshape the organization of work and life across the world. We are focusing our inquiry on these themes. Our research deals with an impulse coming from mathematics, artificial intelligence, and complexity theories. However, while we start from a technology push, we constrain ourselves along the projects with explicit technological avenues, and we link back to explicit market pull demands as we move along. These chapters, hence, provide an overview of the topics we are addressing throughout the dissertation (Lee et al., 2016; Satyanarayanan, 2017).

11.1.1. Overview of the Thesis

Service systems where people work with intelligent systems to provide services to others are becoming increasingly important in today's economy. These systems are increasingly being networked so that many of the individual tasks performed by system users involve communication and information exchanges with an array of other users. In the thesis, I call these systems networked service systems and I focus on issues involved in their design. In particular, I present a framework for understanding and representing networked service system configurations, the implications of that framework for the design of information, control, and coordination methodologies and tools for the networked service system, and approaches and examples of how to apply those methodologies and tools.

More formally, this dissertation explores the design of user-software interfaces in networked service systems. In such systems, an intelligent interface agent, representing local users or service providers, interacts with interface agents belonging to other users or service requestors, to support the user in the performance of a service. The question is not just "How can we design intelligent software interface agents to interact more effectively with people?". It is also "How can we design the interfaces in service systems, where intelligent software agents are intended to represent the user, to ensure helpful and appropriate interaction with other agents?". An appropriate answer must address design issues for both the user interface function and the software interface function, integrating the design of the two. And it must be informed by both service science and artificial intelligence. These goals are the objectives of the work described.

11.2. Fundamentals of Networked Service Systems

1. Definition and Scope

A networked service system can be defined relatively broadly as a set of resources configured to provide some service to the consumer. Networked service systems differ from other types of service systems, such as factory production systems, in that the customer’s involvement in or contact with the system significantly alters the service, as well as the ability of the system to deliver the desired outcome. Examples of service systems include education, healthcare, and hospitality. Networked service systems are a special subset of service systems, in which many consumers interact to create and share value. The best examples are the various communication channels embodied by social media and content-sharing services.

Fig 11 . 1 : Networked Services

The Internet is by far the largest and most widely known example of a networked service system. The only existing large-scale consumer-oriented Internet Service Providers offer an infrastructure for such systems. Other than the ISPs, few organizations offer a general system that enables users to connect, find each other, and share content or experience over the Internet, through specialized ISPs. Design and innovation within this category focus primarily on individual applications, a fact that has contributed to a growing opinion that networked service systems have become something of a commodity.

2. Historical Context

The proliferation of consumer-oriented applications running on the Internet started in the mid to late 1990s, when search engines, email, and news services began to achieve critical mass. This influx of interest reflected in turn the advances in the underlying hardware, software, and communication capabilities – faster personal computing platforms and greater bandwidth at lower cost – that made it feasible for everyone, not just specialists, to build exciting new applications. Today, external users demand more sophisticated capabilities from service systems – increasing flexibility, greater responsiveness to unique needs, and reduced effort – while for internal users of the service system, the objective is to eliminate or at least overcome the vulnerabilities of dependence on these external parties by making the service system more tightly interdependent with the company's core business processes.

11.2.1. Definition and Scope

The development of autonomous systems has sparked renewed interest in how organizations are providing services. The range of research and commercial initiatives that are currently being explored is extensive, including delivery robots, customer service agents, telepresence robots, autonomous guided vehicles, and UAVs used for parcel delivery. Additionally, new opportunities are arising in digital spaces with the widespread adoption of chatbots in social messaging. Over the last ten years, indeed, the services industry has been exposed to increasingly autonomous intelligent systems that are modifying how services are designed and experienced.

Networked services systems are a contemporary and evolving conception of services that enable dynamic, interactive, and value-generating service engagements made possible through the transformative impact of digital information and communication technologies. Networked service systems are increasingly enacting the traditional service-for-service exchange equation in more sophisticated and complex modes from which social and economic values are unfolded in new ways. Herein, this chapter, which adopts a service systems perspective, proceeds to explore the concept of networked services systems in more detail, with emphasis placed on the growing role of new service innovations and intelligent services within networked services systems. Firstly, consideration is given as to the nature and purpose of networked services systems. Secondly, the historical development of networked service systems is explored. Lastly, key elements and components of networked service systems are examined.

To deliver value through the use of networked service systems requires a network of service actors: customers, service employees, service providers, and other actors from whom resources are drawn or through whom value is co-created, co-produced, and received. Networked service systems create value through different forms of service provision – direct, self-service, remote, or assisted service – and in different spaces, including physical and digital environments. Assisting networked service systems are intelligent and operate autonomously or semi-autonomously in physical and digital service environments on behalf of service actors.

11.2.2. Historical Context

Because of their unique properties, networked service systems (NSS) are inherently difficult to temporalize and cut historically, with clear temporal boundaries. However, to better understand and link to existing research, and also define and delimit the notion, we need to outline the larger contexts that the concept stems from, but also previous formulations of similar concepts. Therefore, we will first offer a historical contextualization of the concept, by linking it to several precursors. This is followed by a more detailed discussion of the genealogies of NSS and the different conceptualizations that inspired and framed the development of the present notion. Networked services are not a new phenomenon. The ancient Greeks, for instance, were contemporaries of the first networked service system, in that they commonly exchanged services such as information, entertainment, or assistance with regard to the needs and problems of everyday life for symbolic constellations. Modern societies, however, have until recently mainly focused on synchronized placing and exchanging of material goods in the industrial production and supply chains as the optimal means for achieving symbolic attributes. In contrast, a key property of the infrastructure for autonomy is that it allows and supports the simultaneous provision and consumption of network services within a single transaction. The embedded nature of these transactions weakens their focus on the market, without ever excluding it, and makes them an integral part of everyday life. Along with this shift in focus has come a lowering of transaction costs for network service systems, made possible by advances in intellectual property rights, communications technologies, and tracking-and-tracing technologies. These changes are transforming the way cultural entrepreneurs and creators interact with their target audiences. Likewise, due to the decentralized nature of these systems and the simultaneous use of network services, the development of the service system can be community-driven.

11.2.3. Key Components

A networked service system orchestrates a dynamic ensemble of interactions in which one or more Services and either a Customer or a Custodian are connected to Services or Custodian Interfaces that — preferably in a transparent, efficient, pleasant manner — store, relay, transform, respond, and transduce service information generated by the service interaction. A virtually infinite variety of service information is generated at Interfaces by Services, Customers, and Custodians in a variety of forms for a variety of purposes in the course of service interactions, including socializing (greeting, calling, controlling attention, chatting, wrangling, and otherwise managing social bonds), Customs (facilitating service delivery, travel, and payment; exchanging documents; and passing through borders and checkpoints), Requisitions (requesting services or information), Information Acts (instructing, advising, and reviewing), and Operations (evaluating conditions, ordering services, completing transactions, controlling mechanisms or fasteners, and authorizing actions). This collectively constitutes a collaborative practices infrastructure. Such activities entail the transmission of sensory, symbolic, and machine-readable data expressed in words, gestures, images, actions, signals, and codes, and embody the content and flow of service exchanges, societal relationships, and economic commerce.

It is easy to overlook the fact that service systems are not guided by the command-and-control or governance properly formulated by the theorists and practitioners of hierarchically controlled closed systems such as firms and governments, for example: as unique, essentially social, or open systems. Instead, service systems are supported by a collaborative, self-organizing infrastructure. Nevertheless, Services and Customers or Custodians remain the glue, the anchors that have been, are and will be with us for the foreseeable future. Services, Customers, and Custodians — with a little help from their Interface friends — accomplish the things we call services, requested of or provided by one another, and mediated through socially created, agreed-upon, and maintained service institutions.

11.3. Intelligent Interfaces in Service Systems

Service systems are systems of people, technology, and other resources, who use the Service System for value co-creation. In a service system, direct users can delegate a portion of the work to supporting actors following elements of experience design. The goal of intelligent interfaces is to make the interaction design for service systems which helps service systems achieve their goals with less work from the user and the service system itself. In a Service System Focus approach, intelligent interfaces are discussed in their role as an interface between the supporting actors, which facilitates interaction design.

Intelligent interface components such as intelligent agents and intelligent assisted devices may take on a variety of different formats. For instance, intelligent agent-like interactive voice response helps people navigate through an enterprise's voice messaging system to request services. Intelligent agents and intelligent assisted devices help a user to accomplish most aspects of interaction, coordination, and communication with the wide variety of entities and intelligent interfaces that are part of a service system. Intelligent agents may invoke services through an interaction with a diverse set of underlying conduits such as a knowledge search, a company's knowledgebase or intelligent interface, and person-to-person communication. Some basic interaction models have been proposed to distinguish between different types of intelligent interfaces. The main distinction is based on the extent of autonomy. Intelligent agents have a high level of autonomy assumed to be acting on behalf of the user. At the other extreme, intelligent tools are used like conventional tools but are "smart" in the sense that they have advanced speech, vision, or other capabilities. In the mid-range, user-centered intelligent communicators and interaction supervisors involve the user by actively soliciting their preferences and providing them assistance when necessary or desirable.

11.3.1. Role of Intelligent Interfaces

Intelligent interfaces (IIs) are a core enabler of the autonomous service system vision. They transform the interaction between humans and services by augmenting it with a rich set of physical, contextual, and social abilities. The integration of such capabilities into an interface characterizes IIs in a way that extends work in traditional ICT. Here, an intelligent system is defined as a system that has content - knowledge about some domain - and is reasoning - making decisions based on content, through deliberate capabilities, from an ability space, at a performance level based on task, time, and environment. In IS, the service interface is the boundary at which the service is exchanged, is a convergence of technology and the customer, and is the gatekeeper of the business process. By providing matching services and augmenting the core capabilities of the service system, IIs could redefine the boundary of value creation.

IIs provide task management and transformation capabilities by gathering diverse types of current information including the preferences, needs, location, and state of users and resources; making inferences based on social knowledge and reasoning capabilities; presenting service-related information and requesting input from humans; customizing the user experience; and supporting and motivating users through enhanced usability of service execution. Each task in the service system can be allocated to the most appropriate service agent, resulting in less effort realistically and session times for users and achieving greater service quality by allowing each agent to do what they do best. By augmenting the strong capabilities of automation systems with fine-tuned customizable services, I could ease the transition to automated service delivery for new and sensitive markets while taking projections into account. Adapting the level of automation of an intensive service execution appropriately could maintain service quality in a variety of special situations. Thus, IIs create value for users while augmenting the capabilities of the service system to deliver high-quality service efficiently.

11.3.2. Types of Intelligent Interfaces

The Intelligent Interface is a system of sensors, actuators, and data processing methods permitting automated machines to perceive, decide, and act based on context-awareness in order to work and cooperate better with people in service delivery systems. While the idea of embedding intelligence into service is certainly not a new one, and the idea of creating Intelligent User Interfaces is more than 20 years old, the creation of truly Intelligent Interfaces enabling the automation of joint service delivery performance has gained renewed attention. A wide range of such interfaces has been proposed recently for a wide range of service tasks, with research being conducted to integrate and combine their different capabilities in versatile service robots.

Intelligent interfaces are endowed with sensing, message generation, acting, and cognitive function capabilities, which can be internal, such as based on artificial imagery and mechanical actuation, or external, such as based on remote sensors that provide overlying data usable for the specific function or decision. Based on these properties, we propose differentiating intelligent interfaces based on their sensory data and communicating modalities: Gestural intelligent interfaces allow the realization of a service function jointly by a person and a robot and support ease of task switching. Optical interfaces are limited in their perception of restricted luminous signals and are therefore task-restricted, while communicative robots enable timely communication of high task complexity and support low mental burden. Talking robots have been successfully applied for various tasks in entertainment areas and the early phase of healthcare robots with elderly or impaired persons.

11.3.3. User Interaction Models

The relationship between people and intelligent interfaces resembles in some respects the earlier development of human-computer interaction. Although the concepts are still developing, the user interaction models for intelligent interfaces differ in one important respect: service system design with numerous users and multiple intelligent interfaces tightly coupled into a networked service system. For conventional HCI, consistency mainly refers to the relationship between a user and a specific computer over time. In a typical service system with an intelligent infrastructure, user actions at one intelligent interface influence the intelligent infrastructure and the behavior of nearby intelligent interfaces at or near the same time. Thus, user action consistency and an awareness of other intelligent interfaces modify the design.

The early approaches to intelligent interface user interaction models drew on AI and robotics, with a focus on augmenting the skills and capabilities of users as they performed tasks, often of a complex or long duration. These task-oriented interaction approaches were supplemented by a focus on educational, training, and therapeutic services provided by robots with social interaction and social service robots. These interactions relied heavily on the social relationship between people and robots. The development of sociable robots gradually moved the focus of intelligent interface interaction design to the task and away from the social interaction, consisting of establishing the robots as interlocutors and then labeling the robot's social status. Current tasks performed by autonomous vehicles involve advanced interaction tasks with multiple human users who are the target of service-oriented information, messages, cues, advice, and communication, to reach a beneficial decision or action agreement with the vehicles, contributing to the completion of AD.

11.4. Design Principles for Autonomy

In this chapter, we derive a set of design principles for networked service systems using intelligent interfaces. Users are rarely devoid of control in a successful networked service system. However, the complexity of a design can challenge users' ability to understand, from their perspective, the role of an intelligent interface in the geospatial distribution and management of other networked subsystems. Consequently, these design principles must be user-centric, scalable, and flexible, and enable the design of interoperable systems with intelligent interface components.

1. User-Centric Design

Intelligent interfaces are more likely to be successful in user-centric applications. A user sometimes delegates responsibilities for network decision-making to intelligent components. However, intelligent interfaces cannot be user-centric across all potential use cases, and the role of the user in networked decision-making is defined by the identity and nature of the independent subsystems that are networked by the intelligent interface. A user-centric interface has two attributes: the human user can control the decision logic of the interface and the intermediary roles of the intelligent interface are decided by the user. Assume now that a user-centric intelligent interface has been designed. It bypasses the typical communication infrastructure by enabling wireless peer-to-peer connections between colocated independent elements that have no previous relationship. A user may rely on the intelligent interface to further redistribute those relationships based on previous uses. These types of models can be designed for independence from other intelligent components or sub-networks that come into proximity occasionally or for a longer duration. Initial contact and exchanges may be purely impersonal for logistical or economic reasons.

2. Scalability and Flexibility

In addition to the dimensional and functional topology richness offered by networked service systems with intelligent interfaces, a high degree of scalability must be imported into their designs. For example, to be effective as networking tools, intelligent interfaces must also become small in size, low in cost, and energy-efficient. Unlike the more commonly found model of resource efficiency, the second element of scalability allows the system to become less efficient under stress as the number of devices in a network grows, with a predictable degradation in performance.

11.4.1. User-Centric Design

There are a multitude of designed systems that impose highly prescriptive rules about what can be done and how it can be accessed. Such systems have been designed without recognizing potential use cases which a well-informed user would want to engage in, and/or without understanding the background and knowledge level of the end-users. In designing for autonomy, however, it is valuable to invest in a deep understanding of who the users are, what they are building (and why), what actions they need to undertake, how they would prefer to interact with the networked service system, and how the system can communicate its capabilities. Well-designed intelligent interface elements can help fill in the gaps in user knowledge that persist even after deep user studies and testing with the system. Even with a preference for these interface elements for supporting user engagement, designing a networked service system is not merely a sequential process where a corresponding intelligent interface is chosen for each user-engaged task. If there is a clear delineation between the tasks being performed at the intelligent interface and those being done at the intelligent interface by the end-user, there is still an opportunity to optimize the overall process in a more integrated way. For example, while an end-user may want to rapidly check on the current configuration or status of the service system on their mobile device while on the go, it may be more efficient for that checking to leverage less-capable but increasingly popular embedded elements. However, how the entire task is envisioned and broken up between the intelligent and the designated intelligent interfaces can be influenced by flexibility in both the intelligent and service interface elements. In advanced autonomous service systems, there is a rich functional budget available, much larger than what is feasible for early-stage implementation. With the intelligent interface being such a critical node in user engagement and task motivation, it is also critical to analyze the service system from the behavioral abstraction level into the lower-level instantiation phase of the intelligent interface design. Under the combination of appropriate function and effect, it will be the task of the intelligent interface to optimize the engagement experience.

11.4.2. Scalability and Flexibility

By scalability, we mean that the design should scale with the users' problems, that is, make it easy for new users to interact with systems but push proficient users towards greater productivity. By flexibility, we mean that each interface should support and maintain multiple interactivity styles while allowing digital intelligence to perform an increasing amount of work. Here we talk about both general interfaces and the more special ones associated with creating the functional baseline and a repository of basic functions, algorithms, and resources. Creating social networks that support transactions of aggregate-level data and a marketplace associated with decision support provides both the desire and need for a more standardized, coarse-grained interaction.

Our approach is to allow the initial interface liter, both users and the interface-generated computations and transactions at the design stage must be easy to establish. However, once established, the users and their transactions must be self-managing, with the underlying networking and algorithmic support performing most of the work. Early attempts to do this typically generated a huge amount of unwanted messages and computations. Consequently, most of the early efforts at digital assistants focused on single-user interactive agents, while still allowing the service to be multi-user, although service creation, use, and maintenance were more ad hoc. Speech interfaces rapidly went beyond simple, user-initiated dialog to become the interface for various services, such as travel and appointment booking. Multimodal interfaces expanded both the types of user models that could be sustained over varying application domains and the number of simultaneous users possible. Simultaneously, the interfaces migrating to desktops, PDAs, and embedded devices enabled more complex or simultaneous interactions.

11.4.3. Interoperability Standards

Similar to other industries excited by concepts of distributed ownership and decentralized control-sharing, service robotics development has become more open and less financially centralized. Efforts have developed hybrid vehicle hardware and software platforms and ROS. Pursues service robot open standards with publicly-funded partners, then realizes user functionality by binding interoperable modules. As market demand catches up to developer capability, infrastructure for self-supporting grows–for example, commercializing of consumer bric-a-brac by winning partners and advertising supported local classified markets. Developers can externalize software support costs with fee-based publishing, and/or by developing paid relationships with external sponsors. Future R&D, selling services built on existing platforms developed "under the radar" and passed around folders, could endanger self-supported, (repeated) consumption. For example, quietly implemented helpers using HRC and sophisticated visual modeling.

Interoperability standards can reduce integration costs and simplify developer community participation by increasing the drawing power of dominant service robot platforms - the easier and more exciting it is to build on a platform, the greater the number of offerings for a sponsor, and the more people want to do fun iterations. The goal on the road to autonomy should be modular plug-and-play architectural components with simple but versatile behavioral interfaces. Too many groups have used a successful core technology driving the visual posers behind several commercial hits as a quasi-standard standard, ignoring the duty of a quasi-standard to act as a base of ground for standards shared group responsibility.

11.5. Technological Frameworks

The frameworks outlined in this chapter arise from technological capabilities whose trends have been highlighted above: ubiquitous communication, distributed computation, and intelligent user interfaces. We briefly discuss these foundational services for autonomy, focusing on cloud computing and edge computing services, and natural language processing applications. Together, they function as the technology stack for the Infrastructure for Autonomy paradigm.

1. Cloud Computing in Service Systems

Existing cloud services can be repurposed to play a central role in the design of service systems with intelligent interfaces. Existing technology, skills, and knowledge can be leveraged in innovative ways. Automated natural language understanding, translation, and semantic matching; integrated conversation and dialogue systems; and conversational and multimodal interaction are examples of such capabilities. The advantage of cloud services is that, first, it is easy to understand what kinds of capabilities are available, and cloud companies make it easy to plug these capabilities into new applications. More importantly, the infrastructure that makes it possible to deploy these capabilities for users around the world at scale and at a reasonable cost has already been built and is made accessible at scale. Cloud services will likely continue evolving to provide other capabilities that can support intelligent interfaces.

2. Edge Computing Applications

The edge of the network is becoming a richer computational environment. Dedicated hardware accelerates new or computationally-intensive services such as real-time video analytics and augmented reality. Even if these services are run on the cloud, they need low-latency access to the user - and so are often better run on dedicated edge infrastructure. The edge is also a natural place for other types of middleboxes. These functions can route, summarize, optimize, and cache low-level digital signals that reduce the need for cloud processing and expensive bandwidth.

3. Artificial Intelligence Integration

Like all technologies, the distributed computing and intelligent interface capabilities outlined above evoke fresh opportunities for innovation and novel designs. These technologies have exciting implications for service systems' Intelligent Interfaces. Domain restrictions and characteristic types of interactions help integrate functional building blocks from major cloud and edge services. Well-defined architectures enable the partitioning of computation, language, and vision functions between mobile clients, network middleboxes, and cloud servers. Templates enable conversational experience designs tailored to existing capabilities to bootstrap early applications such as automatic text translation, search and computation, action initiation, and query-answering dialog.

11.5.1. Cloud Computing in Service Systems

Cloud computing is now seen as a powerful infrastructure technology that can enable computer, storage, and communication system capabilities to support autonomous service systems, both in how services are delivered and to what functionary services are delivered. Cloud services are leveraged to enable intelligent services with capabilities and delivery models that were only imagined in the early days of service systems research. We speak to the journey that cloud computing has taken and the trends that service systems engineers are leveraging both to support and to take advantage of enabling capabilities.

Fig 11 . 2 : The Cloud Infrastructure

Taxonomies of Information Technology used in service systems provide an excellent foundation for understanding the basics of how cloud computing fits into the service system landscape. In this chapter, we expand on this work by considering the major service types: Infrastructure as a Service, Platform as a Service, and Software as a Service. The basic enabling and orchestrating technologies are the major fundamentals behind cloud computing and we will discuss these in enough detail to build a platform for sports analytics and other examples in later chapters. Each technology has a limited but growing set of use cases and needs in intelligent service systems and we will close out the section of discrete and hybrid systems. We then consider the details of edge computing and the need for support in the core cloud space.

The mix of IaaS, PaaS, and SaaS environments changes the way intelligent service systems behave from the more familiar world of business process systems. This is particularly noteworthy in the example of the human capability “constructor” function in the delivery of services and is exemplified in the need for new authoritative positions in service systems, such as data and algorithm curator functions. These new roles must work side-by-side with curators and orchestrators of managed service classes in markets of citizen service agents to ensure that the service systems function as intended without unintended glitches or issues.

11.5.2. Edge Computing Applications

Infrastructure for Autonomy: Designing Networked Service Systems with Intelligent Interfaces

Public edge computing resources are becoming increasingly available, with some offering attractive pricing models and approaches for application developer integration. A few examples are: where developers can run applications at the edge of mobile networks, with the advantage of low latency and offloading traffic from core network services; where developers can define and execute some processes in nodes, avoiding latency due to data download from services or allowing users to even bypass providers itself, at the risk of giving up on many services or integrations; where developers can execute applications with high AI workloads with the advantage of low latency; or specialized edge services focused on robotics.

Much of the current edge computing activity is focused on offering developer tools or specialized services in scenarios where edge computing advantages seem to dominate. Nonetheless, the interest in this technology has increased dramatically in many application domains. Edge computing is being used for applications such as systems for smart homes, smart cities, or factories with processing and data analytics modules running in edge devices, filtering the data before sending it to central services; platforms with teleoperated robots capturing 3D video and audio streams in real-time, to then be used in immersive telecommunication and conference services; and fleets mapping their environment and collaborating to process data and optimize logistics; or service systems detectives at the edge capable of detecting important user preference changes, bottlenecks, or detect low-environment risk situations for timely intervention when executing services autonomously.

11.5.3. Artificial Intelligence Integration

In recent years, artificial intelligence (AI) achieved significant advancements. AI methodologies were developed for many challenging tasks, producing results that either exceeded or matched findings produced by humans. The availability of large data sets, the advent of new processing technologies, and the presence of open-source implementations based on popular machine-learning platforms triggered a new wave of innovation. Large service companies applied AI-based products to automate traditional activities carried out by their employees. Is it a result of these businesses being driven by the desire for cost minimization and profit maximization? Or, rather, because of the aspiration for deleveraging, transferring responsibilities, and, ideally, erasing competitive rivals from the service provision landscape? In a way, both. Replacing humans with intelligent systems has enabled high efficiency in the provision of the service while inspiring an ongoing discussion regarding the ethics and consequences of such moves.

With the proliferation of robotics and the growth of the service sector, the use of artificial intelligence for networked systems of robotic service-providing systems has inspired many applied works and thinkers. The justification for such investments is provided by the cost savings and profit maximization achieved through the replacement of robots' physical presence with systems using AI-based solutions.

For any service provider considering the integration of AI within its activities, the level of integration is a key choice. The "autonomy" of the service is at the core of the provider's competitive advantage – and, ideally, differentiating capacity. The service may be delivered either with no contribution of AI, called zero-choice services; or at low levels, defined as human-assisted services: some coordination and control of the service delivery are carried out through the utilization of AI, leaving the final delivery responsibility to the human operator; and at high levels, defined as assisted-public systems: services are delivered on full auto-mode, as the entire process is carried out by the AI solution.

11.6. Network Infrastructure Requirements

A fundamental design goal is to ensure that the physical network infrastructure can support whatever level of bandwidth, low latency, security, and remote data management that current and future applications will require. The design of this underlying network cannot be separated from the design of intelligent interfaces or services. As most users are fairly network naive, these interfaces need to address a whole range of highly technical issues on behalf of the users. For example, issues related to assigning applications and services to the appropriate architectural level dictate the type of wireless communication links used. This discussion also directly relates to the issue of physical versus mobile agents, in the sense that agents need to be appropriately assigned to the appropriate architecture.

1. Bandwidth and Latency Considerations

Bandwidth and latency considerations inform the level of real-time service management possible. Different types of infrastructure may be appropriate for remote command and control, such as telemetry or video for disaster relief, as opposed to supporting multimode virtual service information systems for normal downtown commercial areas. Real estate costs associated with the choice of wired versus wireless will impact interface design, as will physical data transfer rates associated with each of the available technologies.

2. Security Protocols

As infrastructure may include a variety of types, many different security protocols and levels will be needed, just as there are varying policies on usage. These will operate on different layers, from securing each device and interface to ensuring only trusted users have remote access to data management services databases, or are allowed to participate in a group message and ad hoc virtual service formation. Provisions will need to be made for unused data storage capacities on mobile devices to be engaged and for remote management of this mobile memory. In-network devices may also provide security infrastructure services for closest point-to-point relays of interface users, as well as authorization of a logical transport layer for service data streams.

3. Data Management Strategies

As data management will be task-specific, data management strategies will need to be available for retrieval, service agent routing, data format conversion, and real-time transport at various levels of transport latency and data organizational overhead. Information handling at the remote databases will need to account for issues of pictorial versus real-time simulated presence in virtual services, as well as the physical location service matching process needed to enable routing and base transmission services for intelligent interfaces.

11.6.1. Bandwidth and Latency Considerations

To effectively support the deployment of intelligent interfaces in networked service systems, a dedicated infrastructure is required to enable bandwidth and latency at the service level that aligns with the physical systems and digital services being instrumented, autonomous, or interjected. Existing centralized network infrastructures will be increasingly pressured by the need to accommodate massive machine-to-machine communications with growth driven by the demand for real-time services and adaptive interfaces that are becoming part of everyday life. Future networks will need to support a combination of three-dimensional communications ecosystems incorporating satellite and terrestrial nodes, backed by small cells, relays, and edge computers, with network slicing and specific quality-of-service levels that will enable disparate application classes to be serviced to specific requirements. This creates opportunities and challenges for new approaches, optimization techniques, and protocol and operating system abstractions for avionics, automation, and device federation linked by distributed networks across diverse space, air, and ground geographies.

The necessity of achieving low latency at the service level is evident at the application level in the physical services sometimes being provided. In robotics, teleoperation requires round-trip time in the latency range of 20 milliseconds or less to support the smooth operation of flying and wheeled vehicles, which is below latency thresholds for sitting, standing, and walking latency-sensitive conditions. In medical services such as surgery, latency in augments, especially when coupled with a virtual motion model, must be kept well below levels that will compromise the surgical task. Rates for these services are typically based on the protocol of tens of milliseconds, exclusive of faults or network errors, and depend on the aerodynamics and inertia of the robotic and human bodies. In the augmented relay, network round trips will need to be in the flow times less than 30 minutes per every kilometer of distance traveled or 0.05 seconds per kilometer of distance traveled, exclusive of sensor and actuator errors, backhaul and front haul latencies, and predicted navigation speed, which for surgical purposes, will need to operate at 100% sensor, actuator, and communications reliability.

11.6.2. Security Protocols

Autonomous systems often include some form of wireless communication with the home infrastructure. For example, sensors might send information to the home gateway or participatory agents might need to download commands. This leads to two security issues. The first is the risk of eavesdropping or interference by third parties. The communication usually contains privacy-sensitive and/or safety-critical information, which must not be exposed to other people. A more serious risk comes from attackers hijacking the communication channels. They may inject fake commands, corrupting the agent's local state and possibly endangering the agent itself or any human being in its vicinity. This is particularly problematic when agents are mobile, expose their operation to anyone in the local vicinity, and depend on home infrastructure for high-level supervision. In these cases, the home infrastructure cannot rely solely on local, physical trust but must authenticate the commands issued through wireless communication. Moreover, in the case of sensor data, the home infrastructure must authenticate the sensor information, to avoid injection of forged status messages by a malicious agent that acts on behalf of the home infrastructure.

Wireless communications using established standards are sufficient to counteract eavesdropping and interference risks. These protocols authenticate clients upon access to the wireless LAN, using a key that is shared among all clients and access points serving the corporate or home LAN. They further encrypt the transmitted packets using periodically refreshed symmetric keys. The security protocols do not, however, provide per-packet authentication, which is necessary to guarantee that the packets received by a server come from a trusted client. Moreover, the trust relation is symmetric, which implies that an attacker who gains access to the shared key can impersonate a legitimate claim and inject or modify packets at will, although usually for a very limited time. This makes secure communication based on shared keys unsuitable in the case of mobile agents.

11.6.3. Data Management Strategies

The data management strategy for networked SIs is informed by three considerations: the burden of excessive communication overhead, the need for autonomy in local service execution decisions, and the expectation that some local communication still takes place in the service system. Every design attempt to address these considerations will likely require some trade-offs, as violations of the constraints at any point are precursors to service failure on different social, economic, and technical levels. Local overload occurs when a service provider SI must burden the local interface with excessive communication for high-level execution management and local situation awareness. Decisions must still be communicated up to a service SIs boundary for execution management, but on-demand Push or have-good-faith Pull communication does relieve local interface dependencies for situation awareness. A design must balance the cost of signaling up to the network against the risk that local deviations will jeopardize the business operations of the service system.

However, the level of risk that can sustain low-frequency communication will place limitations on local execution autonomy. The risk that a local SI will engage in behaviors that delay service recovery or worsen service system function drives down how much local autonomy can be allowed during contingencies in the local situation. The principle of affordances applies to local communication in that high levels of risk that disrupt business operations will require low levels of autonomy and continuous local communication. Furthermore, the sensitivity of local decision-making to local signals determines the level of continuous local communication that should be required. The higher the sensitivity, the less costly all-continuous monitoring of the local situation will be.

11.7. Challenges and Limitations

When discussing the conceptual development of autonomic service systems, we should address some of the technical and ethical concerns these systems pose. It is important to note that such discussion should be focused on the notions and concepts described in prior sections. We should not be confused by associations with other systems often labeled autonomous. Most of the concepts proposed in this document are practical in the sense they will be implemented. However, we acknowledge the fact some of the concepts may not be implemented in the first models of autonomic systems or there may be optimizations that will be performed before reaching the designations and functions described. Hence, while service systems designed with asymmetric service cycle patterns may be the first to adopt autonomic characteristics, we should not be confused by a limited use of such characteristics with perceiving service contracts as autonomic contracts or servitization as an autonomous principle. Therefore, the goal of this section is to review some of the challenges and limitations of the parameters described, when used to design service systems based on enabling the autonomous performance of service operations.

It is not enough to identify near-end service systems with no front-end physical contact with the customer as autonomic systems because they make limited use of their followers as means of service operations performance. Since the service exchanges are coordinated and deployed by the service provider, either as planned or in response to contingencies, the asymmetry of control and decision-making of the service process is not altered. Therefore, its servitized performance should not be confused with autonomic contracts as they should operate at the service level. As highlighted in previous sections of this document, service contracts that imply service autonomy are not common nor should be with autonomic regulations guiding its corrective interventions to be properly implemented.

11.7.1. Technical Challenges

Our focus in this section is on technical challenges. Many of the ideas for networked service systems with intelligent interfaces, including market-based design, depend on unknown technical capabilities or limits, often involving Herculean feats of engineering, to create something that makes sense in deployment. While much of networking is simply getting it right, in a world where the physical task environment is direct observations or indirect conjectures mapped to the virtual task environment, the union of networked service systems with intelligent interfaces is a whole new ball game! Why? If this is to work reliably, we need a whole new physics to balance task allocation incentives and the properties of the link's latent task link distribution that defines task constraints, all while balancing performance payoffs against the identical observations or conjectures of the task environment. The fundamental question is: what aspects of task environments whereby mapping design is based on conjectures or observations in service allocation can be controlled while the rest are random or unknown leading to pathologies and what ones are ruled out by this best-case?

Establishing a "best" physical model is the first challenge, one that becomes much harder in ill-defining and deliberately incorrect physical models resulting in "surprise" outcomes. However, this is not some combinatorial optimization problem over a search space of primary data. For a deployable system, we need semi-parametric or nonparametric approaches, where the number of misassigned resources, or the use of common flavor generic overfitting models are controlled and where misallocation penalties are modeled. The second big question is: can we create semi-parametric models that produce some of the incredible systemic performance estimates like intuitive scaling distributions as well as the spectacular high-dimensional joint distribution fitting to incompletions of quadratic loss deep imagery models? If yes, this needs to be done in a way that identifies risk to guide design rather than optimize for chance outcomes. If not, are tasks allocable if not degree of allocability? How do we control that if some tasks are not linked to the task model?

11.7.2. Ethical Considerations

Considering the proliferation of data collection and manipulation capabilities, and the explosion of networks of devices that make up the infrastructure for autonomy, the role of intelligent interfaces for augmented, sustained decision-making raises important ethical questions. The interfaces we design and their underlying service infrastructures will make decisions on behalf of end users, supporting or undermining their autonomy. In the name of efficiency, some of these services will occult important dimensions of the decision problem. Their imperfect algorithms will impose burdens on users' time, attention, and cognitive skills, while at the same time overstimulating their implicit biases or their desire for social validation. Consider, for instance, the algorithmically-driven feedback, social or otherwise, that search engines, news outlets, social networks, and dating services deploy in order to optimize engagement, desire fulfillment, and ultimately ad revenue. In doing so, these systems enslave our attention, undermine our sense of reality, disempower us as constituents of democratic societies, and promote irrational or antisocial behavior. There are limits to how far for-profit system operators should be allowed to go before a service becomes detrimental to collective well-being.

The role of intelligent interfaces raises additional issues related to the role of humans in the system loop. While autonomous services operate in the name of the end user, their ultimate authority is undermined by an important truth: execution is ultimately deterministic, but allocation is non-deterministic. The outcome of the interaction is unpredictable in the sense that the preconditions for executing the desired action can change without notice. The aspect of unpredictability reflects the essence of market transactions. The implication is clear: while it is beneficial to offload tasks that are repetitive, preferably sterile, to autonomous services, the historical trend of narrowing down the scope of human decision-making must be reversed at the aggregate level.

11.7.3. Regulatory Issues

Lawmakers and regulatory agencies are exploring the regulatory framework needed to properly address autonomous systems. While it is relatively easy for industries to set safety requirements for large, expensive autonomous machines, it is much harder to impose regulations on mobile devices that the general population will rely upon both in public as well as private environments and to protect these environments from inappropriate uses. Various states have arranged to provide oversight for transportation systems. Industry is starting to set standards in areas where they operate, but there is a lot of debate on whether this is sufficient. For most of these documents, oversight is focused after the development phase. Oversight during the system design phase would likely require new regulatory programs or expanding existing ones.

The primary challenge will be establishing proper protocols for systems that incorporate intelligent digital interfaces. Some suggest that verification should be at a level where it is impractical to explore every possible state. Others suggest that verification for large, intelligent systems should concentrate on the theory of requirements-based testing. Before these automated systems represent a significant part of society, a general regulatory structure similar to laws regulating functions such as automobiles, buildings, and food must be developed. These would provide a framework to ensure that the intelligent systems businesses use, as well as those that governments use, are safe and effective. Taxonomies of these systems must be developed because the risk of failure is directly related to the role a system plays and the physical and digital environments where that role is performed. Regulations have been developed, but they currently do not account for systems that incorporate techniques.

11.8. Future Directions

Infrastructure for Autonomy is positioned to contribute, significantly and uniquely, to the emerging autonomous systems of the future. As decision support systems and intelligent interfaces are integrated into the fabric of society, their quality might be as critical to advancing the autonomy agenda as innovations in robotics or AI. Moreover, the expansion of autonomous systems will significantly accelerate in the upcoming era of Space 2.0, the organization of the next block of Old Space and New Space missions in Mega-Constellations – wherein thousands of small satellites working in concert will provide global Internet, Earth Observation, and Space Traffic Management services. The potentially positive contributions relied upon by government industry and society must be carefully considered and planned against the externalities, risks, and contingencies of failure.

A moment of reflection is in order. In the period of rapid development that we have witnessed so far for the space economy, service innovations enabled by the evolving infrastructure of Low Earth Orbit have often outpaced the knowledge-based infrastructure and their roots in Social Capital and Intellectual Property required to guarantee their good societal impact in the longer term. Support systems at the confluence of technology and society are increasingly needed to help Society manage the changes wrought by high technology in the social fabric and how it operates. Even less developed domains than Space 2.0, such as the promise of the Internet of Things, have experienced negative backwash effects like those described previously, including concerns about Nation-State security hacks, adversarial Agents addled by AI processes, privacy infractions prompting Data Protection Regulations, and loss of personal agency led by issues like social media addiction and Radicalization.

For the next block of space economy missions, infrastructure for autonomy needs also to develop in a preventative manner. It is here that the lessons learned in recent years about the conduct of commercial and state-led space activities are relevant with respect to the design of service and Network Centric Operations and their enabling networked service systems. Rather than relying only on Industry Standards defined in a top-down manner by the early movers and big players, we argue that the Organization of the expected trade routes will be facilitated by the development of systems theory for those service and network-centric operations defining the practices and processes implemented by Autonomous Agents at its socio-technical core functions. Directing these efforts at surfacing potential deviations and deviations from the standard in an early visual manner will inform the planning of system contingencies and redundancies that will assure good operational performance across both routine operations and alarms.

11.8.1. Emerging Technologies

Machine Learning and the Internet of Things are already being integrated into networked service systems through commercial products such as smart thermostats and smoke detectors. Such integrations are likely to result in rapid expansions of both the ability and the use of networked service systems, heralding their readiness for advanced applications such as home-based elder care and automatic fire detection and reporting. Yet there are many other technologies that can also contribute to the evolution of infrastructure for autonomy; furthermore, these other technologies may integrate with Machine Learning and the Internet of Things in ways beyond what is currently realized in collaborative physical products. Indeed, all of the advanced technologies that are emerging in the world today have the potential to influence the future of connected networked service systems. Networked service systems are at the confluence of three branches of advanced technology – automation, connectivity, and data support – and there are innovations in each of these branches that may be brought to bear on the concept of infrastructure for autonomy. Among the many examples of such technologies are advances in Internet connectivity, such as very low power wide area networks, satellite-enabled global Internet, and low-cost short-range wireless mesh approaches. In addition to power, the ability to connect personal sensor platforms to the Internet reliably and reasonably affordably is critical to the emergence of service systems based on the Internet of Things. Short-range wireless personal area networks have enabled such connections for many years, and have spawned entire industries in short-range wireless sensors, wearables, and personal area networks. For many years, mobile telephony was the first and often only practical way to connect sensors deployed at scale. Costs and performance limitations, however, have led many systems toward other implementations, such as cellular modem-less satellite connectivity, radio-based short-range wireless mesh protocols, and Bluetooth sensors with connectionless mobile and hub connectivity.

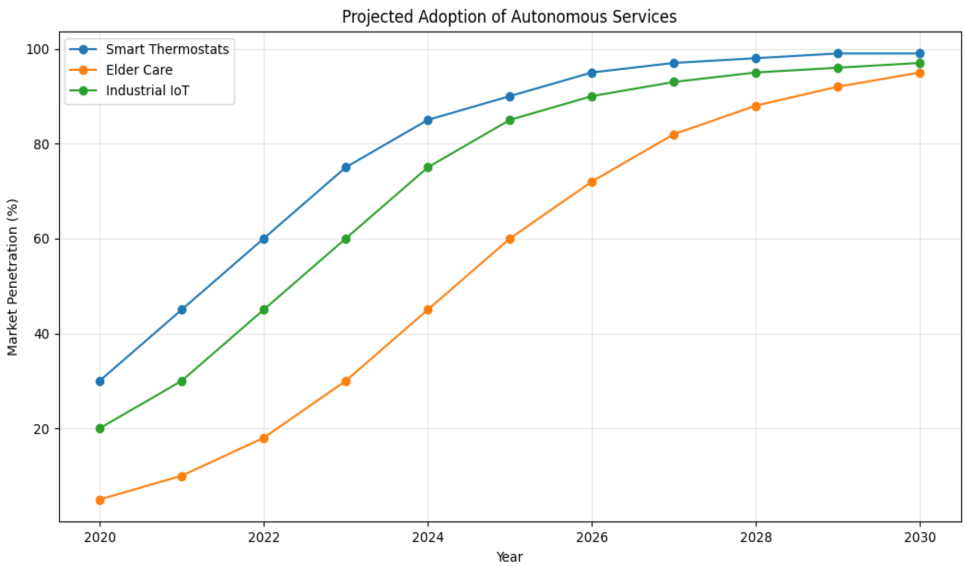

Fig 11 . 3 : Projected Adoption of Autonomous Services

11.8.2. Potential Innovations

Even though there are some products available building on the developed concepts, the entirety of the ideas presented in this work on service design and the service realization chain to support a service system with intelligent interfaces is still to be further explored in research as well as in practical applications. More specifically, in the following, some concrete potential innovations and derivates from our work are being outlined and discussed. The aim is to both inspire researchers and practitioners as well as to trigger more advanced implementations of the concepts presented.

The concept of intelligent interfaces and backend support for such interfaces can also be utilized for augmented reality service applications. By building an augmented reality service application on the nexus between the user and the service, special connectivity features can be provided that enhance service consumption within virtual worlds. Enhancing services within user augmented reality service experiences requires a strong focus on the human user being in all communication with the service system. Ensuring services are also properly rated for both users and service providers – about their service interaction efforts – facilitates that augmented reality applications provide a more service-centered approach.

11.8.3. Research Opportunities

Since the Intelligent Interface concepts and technology are intended as horizontal solutions to enable the creation of new types of intelligent networked service systems, and also considering the strategic importance of this domain within the emerging knowledge economy, we believe that there exist major opportunities for research in this area. Relative to the IId maturation roadmap, the following areas of work can be identified as strategic from a research development viewpoint:

Focus on design for the long tail. The design methodologies, supporting platforms, and tools we envision will aid horizontal innovation but are still lacking today. In particular, we believe that work is necessary for addressing mass customization design, that is, how to support creative designers and craftsmen in the mass production of creative unique objects and services. This is an area identified by flexible, modular design.

Design of modular generic intelligent interfaces. Current work on the development of reusable interface designs, and design templates, is still limited. More work is needed on opening up the design of the user interface, as well as the functionality of intelligent objects and applications. This requires a careful balance between openness and control.

Enabling intelligent interface design automation. The ultimate resources for speeding up the design of intelligent interfaces are, of course, tools that automate part of the design process. These design tools are currently limited. The research challenge, however, is understanding which parts of the design of intelligent interfaces can be automated. In which combination should existing intelligent interface design templates be combined? Which design choices might be suggested to interface designers based on contextual information? For parameterized templates, what approaches can be used to learn parameter values automatically?

11.9. Conclusion

There are many systems in the world that support autonomy in ways that lift burdens from people and contribute positively to physical and mental health. These systems, however, are not the same as systems that 'automate' tasks in ways that marginalize or make redundant the people whose work is displaced in the process. They are instead systems that offer a new kind of infrastructural support to people in their activities and assist but do not do for, people. In doing so, we argue, these systems must necessarily be conversed with, for the objects and services embedded within such infrastructures to be truly useful, usable, and used in materially meaningful ways. These interfaces are, in fact, a new kind of interface, which we call an intelligent interface. Intelligent interfaces are software agents embedded in the networked service systems that support autonomy, helping them deliver the specific support, at the specific time, in the specific way that is needed.

We have identified some design choices that one makes when setting the framework for a service system within which autonomous people might perform their activities, enabled by intelligent interfaces. We have sought to illustrate these choices with design examples and have relied on empirical examples to help us clarify the purpose of the various design moves and mechanisms we have sketched. Our goal, though, has been only to point to the potential of thinking of intelligent interfaces as an intrinsic part of the design of the service system infrastructure intended to support autonomy – as a resource made available, not as a thing that is substituted for, participation in and completion of the activity itself.

11.9.1. Summary of Findings and Insights

In this dissertation, I argue that engineered complex systems need to be designed for service and engagement, as well as for automation. In particular, the development of smart infrastructures that enable people and agents to work together effectively is an important research area in itself and an essential part of the larger enterprise of designing the networks and services that will enable sustainable human communities and the intelligent systems that will support them. Such service systems must be designed as whole systems that embed intelligent interfaces as essential components in the overall system architecture. At the same time, attention must be paid to how the components of the infrastructure are embedded in the larger socio-political-economic context that governs their operation and uses. The insights and findings drawn from this project are presented in three threads: Method, Theory, and Practice.

The project is the first in-depth study of an important and historic class of hybrid human-agent systems called digital collaboratories, which are provisioned by networked material infrastructures, sustain explicit collaboration as a service, and use intelligent interface agents to enhance that collaboration. Because of the longitudinal and embedded nature of the research program in which the project was conducted, these systems were studied at scale, during actual, extended periods of use, as they were built, operationalized, and evolved. This study is the first exploration of how situated engagement, achieved through the mediation of environmental trust, shapes the dialectic between the bottom-up appropriation of resource behavior and the top-down reshaping of that behavior through the emergent design support provided by intelligent environment administrators.

References

Lee J., Bagheri B., Jin C. (2016). Introduction to Cyber Manufacturing Systems. Manufacturing Letters, 8, 11–15. https://doi.org/10.1016/j.mfglet.2016.06.001

Gubbi J., Buyya R., Marusic S., Palaniswami M. (2013). Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. Future Generation Computer Systems, 29(7), 1645–1660. https://doi.org/10.1016/j.future.2013.01.010

Chen L., Miao C., Hao Y., Hossain M.S. (2018). Artificial Intelligence in Networked Systems: A Survey. IEEE Communications Surveys & Tutorials, 20(2), 1011–1038. https://doi.org/10.1109/COMST.2017.2777898

Satyanarayanan M. (2017). The Emergence of Edge Computing. Computer, 50(1), 30–39. https://doi.org/10.1109/MC.2017.9

Dautov R., Zou J., Hu X., Sabihuddin S., Aksu H., Uluagac A.S. (2020). Towards Intelligent Networked Systems: Architectures, Challenges, and Opportunities. IEEE Network, 34(4), 130–137. https://doi.org/10.1109/MNET.011.1900639