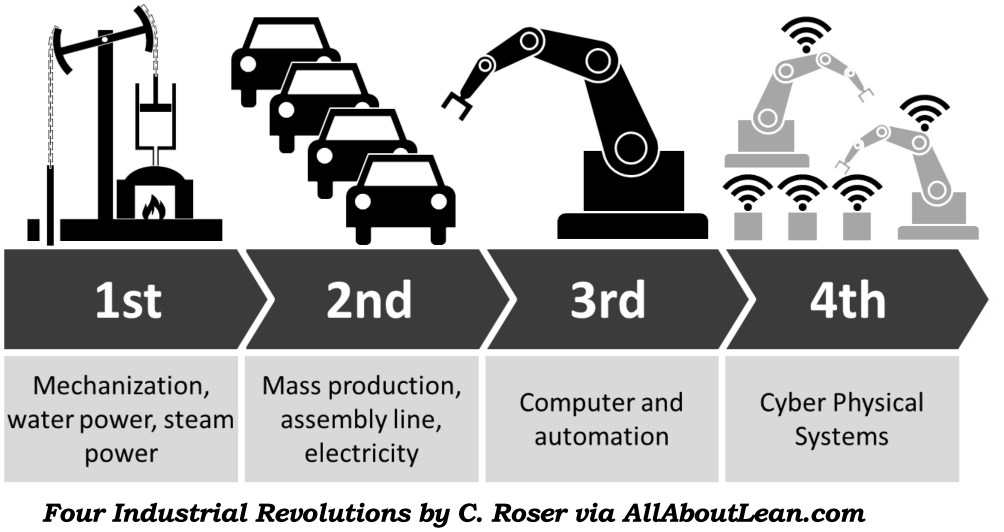
Case Management in Industry 4.0: ACM and IoT

Nathaniel Palmer, WfMC, USA

Introduction

The theme of the *World Economic Forum (WEF)* recently held in Davos, Switzerland was "*Mastering the Fourth Industrial Revolution*" – recognizing the massive transformation currently underway, centered on what has been labeled "Industry 4.0" and addressing the confluence of ubiquitous computing through cloud, big data, and other emergent technologies, but primarily Artificial Intelligence (AI) and the Internet of Things (IoT). The term "Industry 4.0" and the concept of the fourth industrial revolution originates, as illustrated below, from the notion that the industrial revolution represented man's first largescale leverage of mechanical power as a force multiplier in performing work; the steam engine delivered more power than man or horse could then do in any practical quantity. The next major transformation surrounded the shift from steam power to electrical power, greatly expanding the role of worker automation, and the third from which we are now transitioning involving the first round of computers and robotics integration in the workforce. While it can be interesting to put boundaries on each revolution, in reality it has been a long-term *evolution*. Over time man has evolved innovative ways for leveraging technology to perform work in new, more efficient ways.

As was true of previous eras, the 4th revolution represents the continuation of this evolution. Yet is also true that we are amidst an inflexion point in role of technology in work, one that blurs the line between things physical and digital. This "digital disruption" is an exploding phenomenon where software is increasingly a core part of physical devices, the "things" of IoT range from appliances to cars; things which previously had no connection to the digital world, but today are connected through cloud and part of daily interaction. Perhaps most notable at the moment are smartphones and mobile devices, which have nearly overnight have replaced desktops and laptops as our primary means for connecting with the digital world.



Four Industrial Revolutions by C. Roser via AllAboutLean.com

Whether we call it "Digital Disruption" or "Industry 4.0" there is no question we're facing widespread change that, as described by the WEF executive briefing, represents a *transformation on the scale, scope and complexity unlike anything humankind has experienced*. Unlike past revolutions which were largely limited to the factory and industrial environments, this time it will touch virtually every aspect of our work and home environments, ultimately affecting nearly every industry in every country. This transformation is already visible in how *digital* (and notably IoT) has enabled new platforms to aggregate both supply and demand to completely disrupt entire industries and industry structures. Think of the *digital natives* (Uber, Airbnb, et al.) who have created entirely new business models based on delivering goods and services without inventory or infrastructure. Rather, digital natives deliver a platform to connect consumers with sellers of products and services, on-demand and at the tap of a screen.

It is at the customer interface where the impact of IoT will be most pronounced, not the factory floor. That is where customer expectations are radically altered, through persistent, "always on" connections of cloud-linked devices. For the first time in history we are shipping products to customers where we have a persistent, real-time connection. Although you may argue, "but I'm not!" you cannot evade the disruption, as your competitors most likely are already doing it or will do soon. Yet even if they don't, expectations are already changing and the onus is on you to keep pace. This is why it is called "digital disruption" (as opposed to something more friendly like "consumer celebration") because you *shall* be disrupted whether you choose so or not. This also underscores the greatest misunderstanding about Industry 4.0 and IoT; that the value in billions of data-generating devices ("things") is found foremost in the data they generate when, in fact, the real value is delivered through the opportunity for greater customer intimacy. In other words, it is not the data that matters as much, but what you do with it.

BUT WHAT DOES INDUSTRY 4.0 MEAN FOR KNOWLEDGE WORKERS OR CASE MANAGEMENT?

Popular estimates hold that by 2020, in less than five years, the number of connected devices will exceed 40 billion. The result will be incomprehensible volumes of data. The amount of data produced from the beginning of history to the year 2002 is now produced every 10 minutes. By 2020 it will take less than a second. That is an exponential progression in the growth of data, for which there is no reason to anticipate a decline. From monitors and remote sensors, to appliances and vehicles; connected devices are generating meaningful and informative data that would easily overwhelm any human being, but collectively present critical context about processes and the state of operations. Unless it is made actionable, and connected to a responsive loop, the petabytes of data generated by IoT is merely a burden, not a benefit. Realizing competitive advantage comes from connecting the myriad data-generating devices to data-driven processes.

Without the ability to make this data actionable, there is no value. Case management is, at its core, a system of process, rules, and data in support of data-driven processes. It is case management that is best equipped to enable connected devices to add value through either centralized collection of data (ie: millions or billions of data points analyzed to generate meaningful context) or increasingly by pushing intelligence to the devices themselves.

The sensors and even intelligent connected devices that comprise the universe of IoT are the eyes and ears (not the brain), sensing and capturing data from activities and events. This data is then conveyed to actual brains of the system, ultimately involving a set of cloud-based coordination and management capabilities. But increasingly, IoT scenarios may leverage intelligent hubs of automation with local capabilities for storing, transforming or forwarding data, applying a common set of centrally stored (e.g., in the cloud) process and decision logic. This orientation allows for "things" to participate as nodes in the process, including by tying together multiple sensors, and supporting multiple interaction models between both humans and machines.

For example, consider a building automation scenario, where there may be 100s or 1,000s of individual sensors tracking temperature, lights, foot traffic and other discrete matters. At various physical points, such as individual floors or wings, there would be intelligent hubs or local gateways, collecting thousands or millions of data points, and making decisions based on rules and data patterns, about whether to turn lights off, raise or lower temperatures. There would be a central controller that not only connects the gateways, but provides core logic and analytical reporting. For example, an edge-point sensor may not have any intelligence to make a decision, but can provide input to the primary intelligent hub, which itself can apply business rules around event data to decide what action should be taken, such as alerting a building supervisor or escalating to a first responder.

By pushing intelligence into the edge-points, and "smart enabling" existing infrastructure, the combination of IoT and case management allows process automation to be adaptive and data-driven, in turn, expanding the envelope of what can be automated. For example, in the building automation scenario just described, a smaller version of the intelligent hub can pull data from multiple sensors and make a more complex evaluation of localized business events, and either deliver an automated response locally or communicate the context to the central intelligence hub.

CASE MANAGEMENT AND IOT DEMANDS A MODERN BPMS

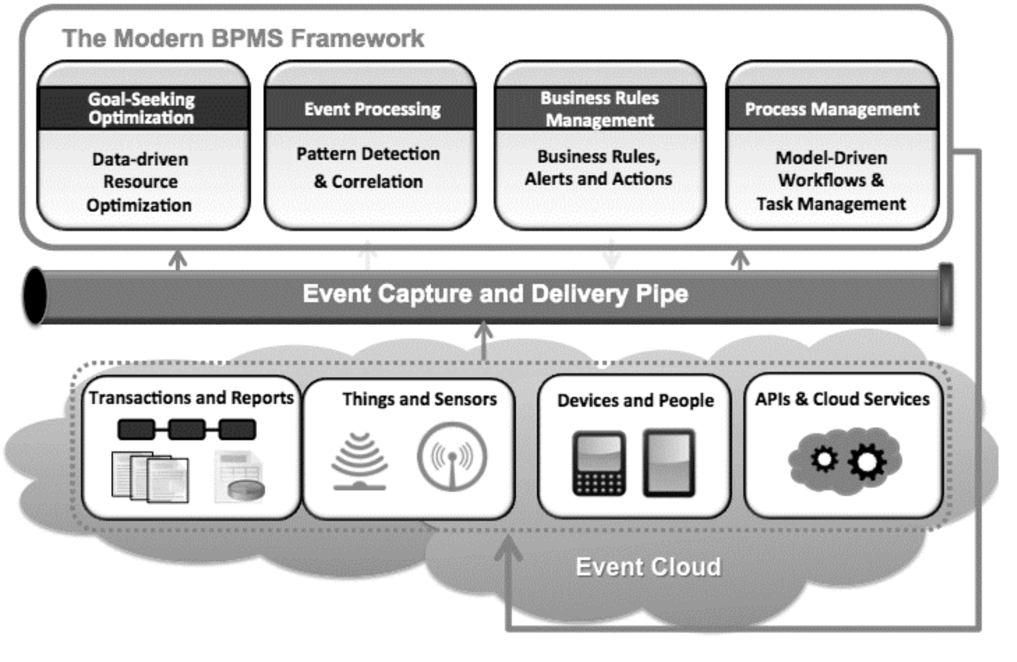
Although case management has been with us for many years, contemporary case management applications, what were refer to today as ACM or Adaptive Case Management, are almost invariably built on the foundation of a BPMS (Business Process Management System). Central to the BPMS is the process model, which defines the sequencing of activities ("sequence flow") and the passing of control from node to node ("control flow") as the process is executed. This is how BPM differs from the basic workflow found in early case management systems that were focused simply on assembling data, rather than actions driven by that data. In contrast, BPM systems manage the process from beginning to end, across multiple participants or "actors" within the process, which may be people, systems, or sensors.

This notion led to a redefinition of BPM in the recent past, in recognition of the shift from basic process automation, to a data-driven model of managing business activity flows, which span one or more systems and where the BPMS coordinates both tasks (work performed by humans) and ***automated activities***.

As the price of localized computing continues to drop, the ability to push more intelligence into the edge-points becomes cheaper and easier, yet doing so hastens the need for process and rules to coordinate the activities locally. These sensors are built typically for a particular purpose and, as stand-alone devices, the data they generate will be of little value to the business apart from alerting to the occurrence of predefined events, such as those in the building automation scenario. Adding intelligence to systems in which they participate does not require better sensors, but provides the ability to identify and define event patterns based on the data they generate. This then allows local execution of automated activities based on the rules and process managed by the BPMS.

This type of "intelligent system" presents a shift from procedural automation, where responses are tightly-coupled to specific sensor data thresholds (e.g., a SCADA system[1] which shuts down a motor when a temperature threshold is breached) to goal-driven models where processes are defined in terms of specific milestones and outcomes (goals) and the combination of rules and data are used to dynamically define the optimal process to achieve them. Such a case is the admission of a patient for medical treatment, involving input from a host of sources and a large volume of system-generated data (even if manually transcribed by a human being) the majority of data involved originates from a device, machine or system. Successfully completing the process requires a "super user" who, for example, knows not only the medical protocols to make a successful diagnosis, but also the system protocols defining where and how to enter and access the appropriate information.

Alternatively, a modern case management system (one built on a platform of BPM) can far more effectively coordinate the streams of data defining not only the circumstances of the patient, but also data defining the various resources involved in delivering care to all patients. This scenario illustrates both why BPMS is the ideal platform to leverage IoT and how BPM as a platform has necessarily evolved. We have developed a new high-level framework for illustrating the modern BPMS platform, illustrated in the model below.



The Modern BPMS Framework (source: BPM.com)

[1] **SCADA** (Supervisory Control And Data Acquisition) is a system for remote monitoring and control that operates with coded signals over communication channels (using typically one communication channel per remote station). https://en.wikipedia.org/wiki/**SCADA**

In the model on the preceding page, at the bottom you see the various access points we refer to as the "**event cloud**," which pulls in data from a broad set of sources, ranging from transactional systems, to sensors and "things," to human users on various devices, plus others programmatic interfaces to APIs and external cloud services. The event cloud is not a physical part of the BPMS platform per se, but the edge points. Connecting these to the "brains" of the system is integration middleware serving as an **event capture and delivery** "pipe" for receiving and delivering event data. Data from the various sources, both internal and external as well as structured and unstructured, is standardized for event processing.

Event processing (also called "Complex Event Process" or "CEP") is applied to all data and business events to generate context based on patterns and relationships among events. In other words, event processing identifies what matters from a stream of events that may not otherwise be relevant as individual events. Looking at this through the lens of IoT, it should be immediately clear why this is so important.

As *James Bond* author Ian Fleming famously wrote, "Once is happenstance. Twice is coincidence. Three times is enemy action." The frequency and temporal nature (i.e., timing) of events is critical to the meaning of events in aggregate. Without event processing, a BPMS or BPM engine cannot distinguish between one or 1,000 events, because each is treated as a single input and would be quickly overwhelmed in an IoT context.

Leveraging event processing, however, allows meaning to be captured from the event data generated by sensors and things, as well as other systems and even human interaction. An illustration of this may be found in the patient care scenario. There may be business rules in place to take action such as increasing staffing for the next shift if the number of patients reaches a given threshold, or if the care involved mandates that more time will be required than originally scheduled. Yet by the time these circumstances are humanly visible, it may be too late. However, there are various events that may occur below the threshold that triggers corrective action that collectively foresees future problems.

For this reason, both event processing and **Business Rules Management** capabilities are required. Rules allow for goal-driven systems to determine the sequence of a process based on current context. For example, a BPM system can examine appropriate business rules and other defined policies against the current status of a process or activity to determine what step should occur next and what information is required. This is, in fact, a process, yet it is based on a diagnostic procedure that likely involves applying a combination of policies, procedures, other rules together with the judgment of healthcare workers. Information discovered in one step can drastically alter the next set of steps, and in the same way, a change in a patient's status may completely alter the process flow in other ways.

The circumstances that define the patient's status are represented within business rules. The sequence of steps to take, in particular, which step to do next; notably who should do it and whether or not it has already been done, is managed by the **Process Management** component of the BPMS. The process management engine manages the control flow and sequence flow of the process, and generates tasks and activities, or specifically initiates these, based on the parameters defined in the process model.

Event processing, business rules, and process management can be successfully applied to goal-driven scenarios involving the massive volumes of data generated through IoT, allowing real-time adaptation based on event patterns and predefined policies and procedures. Yet by widening the aperture of addressable data with IoT, the number of possible permutations and considerations can quickly overwhelm the ability to make *optimal* decisions. As stated earlier, the volume of data generated by sensors and things is growing incalculably high. To truly leverage data-driven resource optimization factoring all the event data in IoT scenarios requires a fourth component: **Goal-Seeking Optimization.**

The value of IoT is not simply the data generated, but how that data can be made actionable. With sensors and other data-generating devices, we can have a real-time or even temporal understanding of resources and constraints. The ability to optimally organize these is the core advantage of IoT. For example, imagine delivering 1,000 unique shipments to 300 customers with 15 trucks and three warehouses in the fastest period of time. This goal is a lot easier to meet when all of these actors can communicate their circumstances, not just once but over time as circumstances change and externalities such as traffic and other factors are presented. Having access to this data is critical, but it is of little value without an algorithmic way to solve complex scheduling and resource optimization problems.

Is this case management? Yes, it is ACM – specifically it is supporting adaptive, data-driven process by empowering knowledge workers to know in real-time what is happening at the edge points, and to take actions through the combination of rule-driven guidance and their own know-how. It is not a traditionally-automated system but *intelligent automation*, the hallmark of Industry 4.0 where technology doesn't merely replace human decision-making but extends the reach of the knowledge worker; making IoT data actionable. In the above scenario the logistics analyst plays a role similar to the head nurse scheduling patients and resources, or insurance claims analyst, all of whom apply their own domain knowledge yet work with "guard rails" of policies and rules, facilitated by a system which proactively manages and guides their work, freeing them from the burden of tracking data, as well as the often-impossible task (for a human brain) of calculating optimal schedules and resource combinations.

What about the truck drivers? Well, this is Industry 4.0 we're talking about, so of course, they're drones coordinated by the BPMS (offered only half in jest).

AI as an ACM Interface

If there is any single concept of Industry 4.0 that is the subject of more speculation and debate than IoT and automation, it is AI. Yet lost in the consternation over machines run amok ala *HAL 9000* is the more realistic and powerful notion of **AI as an interface**. Specifically, think of user interaction with business processes as conversations with "bots" rather than traditional human tasks delivered through a worklist metaphor.

Here's what it might look like… A new task has emerged which requires your attention. An alert is sent to you via the medium that makes the most sense for this task; let's keep simple for the moment and simply say via email. The bot, let's call her "Amy," parses and summarizes the task, and asks for response. In one case, it may be an exception that requires approval and in the summary, Amy includes the context from the decision logic (e.g., business rules) that outline why this is, in fact, an exception to standard, straight-through processing.

In another scenario, Amy may ask you to review a document, then either approve or deny specific circumstances. In the very near future (albeit already possible today) another bot will have already read the document and passed the summary along to Amy complete with rule-driven conclusions and recommended actions, such as potentially-suspect information requiring further examination. In both cases, you need only respond in a natural language statement as you would any assistant, which may be as simply as "approved" yet more likely you will actually 'converse' with Amy such as requesting additional documentation, escalating the work to another queue, or reassigning the work to a colleague.

Here is how this might play out in an expanded use case. An application is received and decision logic is applied to verify completeness and eligibility. The case folder for the application contains all the relevant details (e.g., the application itself and associated application metadata) including the results of the decision logic applied – the rules that establish whether or not the application meets the criteria for approval as-is or (more likely) whether additional information is required. Up to this point it is largely an automated process; the application is received, processed, reviewed and an initial assessment can be rendered. Yet this complete set of information is not necessarily only for basic interactions; it needs to be managed within the case, but most likely is more cumbersome to present all at once.

In other words; show just me just what I need to know and I will decide what information I need to render a decision, as long as you can assure me it has been read, validated, and what judgments I need to make.

Also note that we haven't yet stated what *kind* of 'application' and it largely doesn't matter, because up to this point, the same basic steps would be required (albeit with different rules and decision logic) whether it was for a mortgage, credit card, business loan, health insurance, or another type of application. In any of these cases, the validation of the completeness of information and general determination of acceptability should already be defined within the process and rules, and this is managed by the underlying BPMS. Yet the interaction is managed by a bot or "cog" which can communicate the salient points to me conversationally, and extract from my response the discrete data points needed by the BPMS. The interface should not prescribe how work is performed, nor should the interface be tightly coupled to the task itself.

AI can bridge the gap by interacting with the work management system, in the same manner in which a human assistant can intermediate between the busy executive who needs most often just details, as well as the ability to ask questions to further drill down, or to be told when something requires their undivided attention. This highlights a critical principal for both the future of software, and perhaps more notably the charter for ACM: *to re-envision the structure of the task to be not a single, discrete unit of work, and to remove the distinction between what supports a task and the task itself*.

Let's revisit the original assumption that this is all done via email. Certainly some tasks can be performed via voice commands, just as the currently common example of a bank's fraud detection system by calling a customer directly to verify the legitimacy of credit card transactions. Not all work can be done, however, via voice prompt, text or email, or through an interface of a connected car or wearable device. Many tasks can be thus done and the ability to have work truly follow the worker, for both convenience and expediency, underscores the value of *separating* how work is performed from the work itself.

Do we have to interact with AI on a first-name basis and why "Amy?" Yes, most likely we will interact with bots as people, calling them by name and often without clear distinction whether they are human or cog. Amy, or Amy Ingram (get it? "AI") is the "AI-powered personal assistant for scheduling meetings," which I have recently been beta testing for x.ai – a startup focused (currently) on just that one function. Amy passes the Turing test and even has her own LinkedIn profile. Unlike

Siri, Amy does just one thing (schedules meetings) and does it very well, more pleasantly than any other meeting scheduler I can recall. She's not big on small talk, but she can parse the conversation to extract the details of a potential meeting by being copied on email, reaching out to all parties whenever the topic of a meeting occurs. It (she) is not a standalone system but an active participant, yet the calendar, the email, the function of work management remains in the systems of record. Just as it would work as a BPM interface, the AI is merely an abstraction of the underlying applications. It leaves the core systems intact but greatly simplifies the interaction as conversations, proactively chasing down the other participants and asking qualifying questions until the work is successfully completed, based on criteria defined within configurable rules.

THE INTERFACE IS *THINGS*, NOT A THING: WHY THE INTERFACE MATTERS MORE WITH IOT

As disruptive as smartphones and the emergence of mobile platforms have been to software design, the impact of IoT and Industry 4.0 will inevitably be far greater and happen far faster, on a scale of many orders of magnitude more than with mobility alone. With mobility, there has been essentially a single new digital interface, or form factor content, but the mobile app has been replaced by full-sized application UI as the primary access point. IoT and the broader spectrum of Industry 4.0 represents the emergence of many new access points and patterns of interaction, from intelligent connected devices to AI cogs, to connected cars and social robots, to other new devices yet to hit the market but soon to become common within our homes and businesses.

It is not only conceivable, but inevitable, that we are going to see a massive dematerialization of the user interface to our core systems. Industry 4.0 requires rethinking what the task look like, to remove the distinction between what supports a task and the task itself. This is literally what ACM was made for; to allow knowledge workers to work in the manner and environments that suit them best, at the moment that work is performed.

And there may not be an app for that; this may, in fact, be the beginning of the end of the smartphone. Consider the prediction made by Ericsson's *ConsumerLab* late last year – that within five years, smartphones will be obsolete. Sound ridiculous? Given the growth of smartphones over the last decade, it may seem far-fetched. Yet their reasoning is sound, and is based, not on speculation, but their own consumers' feedback. Half of the study's respondents stated that they expect to be able to interact with objects without the need for a phone or tablet because phones and tablets are merely clumsy intermediaries. They have been convenient in the absence of other "things" but increasingly, we will interact with specialized devices and the more fit-for-purpose interfaces they offer. Our systems need to be able to anticipate that the interface will no longer be represented by a single thing, but by many things. This is the promise of ACM in the age of Industry 4.0 – a challenge of adaptability that ACM is best equipped to answer.