Predictions for 2025

RFG Perspective: 2025 will be a year of optimism and hope that the destruction that the world has endured through 2024 is ending – domestically, economically, financially, and geopolitically. Inflation is abating, but prices and interest rates will remain higher than desired, and the potential for no-growth or business contraction remains a possibility domestically and internationally. The incoming Trump administration is commiting to bring an end to war in Ukraine and the Israeli-Gaza war, and alieviate the concerns for a China-Taiwan war and/or seige. It is likely that its success may take much longer than anticipated. The Miiddle East is in flux – with the expectation that Iran (and its remaining allies) will fall. Additionally, we remain in the age of scarcity, which will continue to affect data center systems, various components, and the energy to power it all.

The traditional assumptions that were true prior to COVID-19 – such as cheap capital, unconstrained growth, low-cost energy, reliable power sources, unrestricted water supplies, rapid access to new equipment, just-in-time delivery, and an economy where products are disposable and replaceable easily – is totally a thing of the past. The new paradigm entails limited growth, energy constraints (at higher prices), potential power shortages and outages, water supply constraints, supply chain constraints, labor shortages, and the need for a reusable, circular economy. All this, in turn, necessitates that businesses modify their business plans to address efficiency and sustainability. “Being Green” is no longer just an abstract commitment from CEOs: it is flowing into business plans (not just greenwashing) and becoming a requirement for IT executives.

AI, and in particular generative AI (genAI), has evolved from being a useful technology to a business demand by CEOs and Boards of Directors, and a concern of politicians and regulators. There are various types of AI. For simplicity, one can categorize AI in terms of system processing analytics, gaming analytics, and genAI. GenAI is a interactive productivity tool that primarily uses large language models (LLMs) to assimilate existing data points (tokens) in order to deliver insights, analyses, explanations, and recommendations. GenAI just celebrated its second anniversary and, while it has captured the imagination of executives, the models and applications still have flaws that need to be worked out. Nonetheless they are being adopted and will be come core to many business applications.

The digital transformations of corporate business models will add to the overall distress and uncertainty, as companies and individuals strive to determine how to survive and prosper in the new world environment. To address these challenges, business and IT executives must re-examine their business models, culture, people, processes and supply chains – and all of the impacts caused by external pressures.

Executives will also need to resolve staffing and talent challenges through collaborative, interactive, and personalized approaches with the goals of improving productivity, diminishing employee turnover, and addressing skills limitations. There will be a need for process changes that leverage AI, analytics, IoT technology and remote sensing, mass personalization, mobility and sustainability. These technologies will connect employees throughout the organization wherever they may be at any moment in time. The same approach will be used to enhance an organization’s connectivity with business partners, customers, and supply chains.

46 Kent Hills Lane, Wilton, CT 06897; (203) 429-8951;
http://www.rfgonline.com/; Contact: inquiry@rfgonline.com

In conjunction with this, IT executives must address data integrity, privacy and service-level concerns that impact business outcomes, productivity, revenues and cybersecurity – all with the goal of building more confidence in IT and the organization as a whole.

To achieve these objectives, IT executives must increase their focus on analytics, automation, componentization, containerization, AI, GenAI, environmental and operational efficiencies, orchestration, privacy, and security. By focusing on these priorities, IT will be able to deliver more and better offerings faster – and at a lower cost – while protecting the business from cybersecurity attacks and vulnerabilities.

The year 2024 proved to be a difficult year globally. Instead of one war, there were two with the prospect of a third. The continued rise in prices and sporadic shortages caused business executives and government agencies to find ways to cope with these not-so-temporary changes. As we move into 2025, people are uncertain as to whether these challenges will be with us for an extended period, or if the changes and rightward shift in governments in many of the developed country governments will turn things around.

In the United States, enterprises are hoping the Trump administration will curb inflation without a recession happening, and that there will be a period of growth across all industry sectors. Outside of the U.S., there is a glimmer of hope for change but turning the ships around will not be an easy task. It is likely most governments will fail to truly produce change. The prospect that wars, conflicts, and disruptions could arise in hotspots around the world is driving business executives everywhere to be cautious about growth and worry about potential negative impacts to global commerce.

On the technology side of business, the various genAI models have generated tremendous excitement in corporate board rooms and C-suites. Visions of new applications that drive new revenue streams and productivity gains are dancing in their heads. There are already examples of success but unfortunately, reality for many is not living up to the hype – but IT will be given the directive to charge ahead and spend on genAI solutions. Given the nascency of various models and the costs associated with genAI operations, targets will be tough to attain, and financial viability will take time to work out. IT executives should expect more failures than successes, particularly as they start their AI journeys, and need to guild businesses as to what reasonable returns can be realized, how, and when.

Meanwhile, cybersecurity attacks will continue unabated, forcing businesses and governments to re-examine their business culture, their processes, and the degree to which they are safeguarding privacy and security. All of this means that executives must invest in transforming their businesses while they still struggle with staff, skill, and supply shortages. It will be a tall order: executives must simultaneously drive process improvements to remain competitive, contain costs, enhance compliance, minimize risks, ensure privacy, and improve resource utilization – all at once.

46 Kent Hills Lane, Wilton, CT 06897; (203) 429-8951;
http://www.rfgonline.com/; Contact: inquiry@rfgonline.com

Major Trends for 2025

Below are six areas that we expect will be focal points in 2025:

AI and GenAI –Most IT organizations will struggle to successfully implement genAI applications. Whether done in the cloud, colocation site, or on-premises, these projects require coordination amongst the IT development and operations teams as well as facilities and security staff, at a minimum. Careful coordination can eliminate or minimize the bottlenecks. The large language models (LLMs) and the processing power required to support genAI training systems can be very expensive – potentially thousands of CPUs and GPUs and petabytes of storage. Many firms will look to small language models (SLMs) for training models to reduce costs and speed up the time to implementation. Not only is the acquisition cost high but the power and space required is also exponentially substantial in comparison to general workloads and even compute-intensive ones, and for many firms, requirements will exceed the existing data center or colocation space and power capabilities.

To support new genAI applications rack density is increasing from less than 20 kWh to more than 50 kWh and data center size is expanding up to 400 mW with campuses reaching one gigawatt. Finding sufficient redundant power sources for these demanding data centers will be a challenge in 2025. Enterprises hope to address this by shifting site locations to areas where power is not max'ed out or by experimenting with nuclear-powered energy sources. The shift to greater rack density will also force the acceptance of liquid cooling solutions into the data centers. In that there are no standards in this area yet, IT executives should consider its early adoption as potential throwaways.

AI in some form will be a required component for nearly all new solutions that require some form of analytics. Many organizations are in the early stages of AI adoption, integrating it within a single function. Most organizations are looking to incorporate AI to a moderate or significant extent in demand forecasting, HR help desks, brand and trade promotions, and inventory and order management. As is already evident, AI, genAI, and ML software will prove to be highly valuable components for most new applications. This extends to customers' applications development processes, IT operations (in the form of AIOps), and OT (Operations Technology) production tools, and specific business-case scenarios to deliver enhanced, intelligent, and integrated service sets. The top AIOps areas will be automated network operations, capacity planning, energy optimization, predictive maintenance, security incident detection and response, and workload resource management. Development teams will also experiment with various genAI code development and analysis models.

In addition to the internally developed genAI solutions, companies will implement multiple external genAI solutions and tools. IT executives should expect that some of these applications will be invoked by staff without corporate approval. Therefore, IT governance, privacy and security groups need to be alert to all external applications (including those invoked with corporate consent) and be prepared to minimize the risk exposures.

46 Kent Hills Lane, Wilton, CT 06897; (203) 429-8951;
http://www.rfgonline.com/; Contact: inquiry@rfgonline.com

Efficiency, ESG, and Sustainability – While most F500 companies have made commitments to reduce greenhouse gases, with the aim of achieving net-zero emissions by the year 2050 or sooner, there is now a shift to back away from these commitments. Fewer than 20 percent of CIOs and IT teams have been asked to directly contribute to these efforts, industry surveys reveal. Specific guidance to-date is mostly lacking and, given the new direction from the Trump administration and its impact on corporate executives, RFG expects to see no change in the guidance or any increased sustainability push by top management.

Few IT executives know where to begin to implement a sustainability/efficiency process or what the key metrics should be. Moreover, these initiatives will be exceptionally difficult to bring to fruition without cross-enterprise alignment and agreement on targets, goals, and the necessary funding. [*For those interested in understanding the data center sustainability metrics and how to achieve them, RFG can help. It is an area that RFG has specialized in since 2004.*]

Regarding the age of scarcity, IT executives can expect pockets of scarcity of energy and/or water in many locations around the world, especially during extreme weather conditions. The addition of genAI solutions will exacerbate the problem, by adding up to 20 percent more in power requirements for new applications. Not only will IT and Facilities executives need to reduce (or at least contain) their carbon emissions by 2030, but they will also need to find ways to become more efficient and more sustainable. The last time that enterprises were incented to improve data-center efficiencies was more than a decade ago – and, sadly, much of that institutional knowledge is now gone. To be effective IT organizations may have to implement data center modernization projects in conjunction with (and possibly as a prerequisite to) colocation or on-premises genAI projects.

While the push for efficiency and sustainability is optional in most countries, the revised Energy Efficiency Directive (EU) 2023/1791, which was adopted by the European Parliament and the Council is now in effect. The new EED establishes an EU legally-binding target to reduce the EU's total energy consumption by 11.7% by 2030 (relative to the 2020 reference scenario). It sets annualized targets and requires each member state to transpose most of the different elements in the directive into national law within two years.

Meanwhile, California passed the Climate Corporate Data Accountability Act (SB 253) and the Greenhouse Gases: Climate-Related Financial Risk Act (SB 261). The CCDAA requires large businesses to disclose both their direct and indirect greenhouse gas (GHG) emissions. That means data center operators will have to report the emissions from their facilities – i.e., accounting for the carbon footprint of all of their servers, storage, networking, UPS, HVAC, and other data center equipment. In addition, they will also have to report emissions from their supply chains, transportation, and business activities, which will be no mean feat. IT executives can expect this to cause massive confusion and require significant reworking in how data center metrics are measured. And, as California does, so does the rest of the United States in time.

Not to be outdone, the U.S. federal government passed the Federal Data Center Enhancement Act of 2023. This has the General Services Administration (GSA) establishing minimum requirements for new data center – including but not limited to availability, sustainability, uptime, and security components.

Given all these new directives, IT executives will now need to demonstrate progress and provide plans towards the reduction of energy usage in their clouds, colocation sites, and corporate-maintained data centers. The good news in this: in the EU and elsewhere, there are significant actions for IT and facility executives to take that can reduce their carbon emissions and power usage by 50 percent or more compared with current levels. However, it will require cooperation between multiple business units, all working together to understand their current baseline metrics and create a fundable business case that can achieve corporate-wide results for improved efficiencies and sustainability purposes. This is another area where RFG has found that movement to an OpEx model and an expansion of the lower cost hyperconverged infrastructure (HCI) will provide IT executives the most flexibility in addressing efficiency and sustainability near- and long-term.

Business Transformation Journey – The shift to the digital economy combined with the work-from-anywhere/anytime environment, demand for genAI applications, and the scarcity of resources continues to force organizational business transformations. This is likely to be the norm going forward. Business processes are, indeed, being reshaped.

Work processes will be in flux in 2025 as companies attempt to interact via a standard set of genAI, social media and software tools. However, security, privacy, and data-governance guardrails must be constructed to minimize the exposures created by the evolving work environment. Bottom line: IT executives must work more closely with line of business (LoB) executives to ensure that funds and resources are being allocated more effectively than they are today. Alignment requires a three- to five-year outlook, planning, and strategy.

Software that ensures new functionality will require automated oversight, usually by deploying AI/ML software for event discovery and rapid identification of IT issues and use of genAI for predictive analysis, code development, and draft recommendations. Firms will automate more of their business processes and operations, perform process re-engineering, and re-evaluate supply chains and staffing. Lastly, companies will tackle the online customer experience by adding genAI solutions that support natural language because users prefer working with enterprises that simplify the process of doing business with them, making interactions more natural and frictionless.

Cloud/Edge/IoT – The network/server/storage topography continues to evolve, even more so now with the arrival of genAI applications. Enterprises will be more selective in their choices of where to place applications and data – with the aim of having the work being executed as close to the physical work site as possible. The secondary location parameters will be complexity, "fit for purpose," and management. One necessary result of this will

be a slowdown of applications and data moving to public clouds and greater use of colocation facilities to reduce cost and latency. These considerations will make it difficult for IT executives and cloud architects to construct application and data architectures that can survive long-term without sparking major outages or failures. 2025 will be a year in which enterprises must address their cloud management and BC/DR requirements, as the frequency of outages by cloud service providers (CSPs) is forcing business and IT executives to recognize that clouds are not always available and that CSPs do not have all the availability answers – yet.

New methods of abstraction, automation, software orchestration, and portability are emerging to address these growing issues. But many companies still find themselves locked into legacy on-premises and cloud solutions with "islands of data" that remain disconnected to other key systems, including on-prem and off-prem systems. Having these data islands, or data silos, makes the productive use of business analytics far more difficult, if not ineffective or unachievable, for IT and for the business it supports. Many HR, marketing and engineering executives are looking for genAI applications to solve this issue in 2025.

Cloud affinity and lack of portability means that architects must determine the best target Cloud/edge solutions for each workload – and do so early in the design/development (DevOps) cycle. While more workload migrations to clouds, edge, and IoT are planned for 2025, the business impacts and business results will likely be mixed unless the design goals are clear – and the impact of applying these technologies is understood. In short, determining the true impact of these technologies on business development will depend on addressing the full spectrum of cross-organizational issues within and across the extended business environment including subsidiaries, partners, and the supply chain.

Cybersecurity and Regulatory Compliance – The number, sophistication, and variety of cyberattacks is expected to increase again in 2025. The use of analytics, along with artificial intelligence (AI), machine learning (ML), and genAI software, will improve real-time analysis and reduce the risks. While technologies for fighting cybersecurity attacks will improve, RFG believes that cybersecurity breaches will remain a major challenge that IT executives must address in 2024 for the following reasons:

- The security guardrails are still too lax to prevent security lapses.
- Human error (e.g., via software design, customer service, and user error) creates security gaps that can be exploited in cyber-attacks and ransomware.
- Third-party software contains vulnerabilities that could cause inaccurate or unexpected damage to the business.
- State actors – and non-state actors – will aggressively add genAI tools to push to find the weakest links in an organization's security and data-protection software.
- Ransomware attacks have proven to be a successful business – for the attackers, gaining them substantial – and unearned – funds.
- Regulatory agencies continue to write new or revised laws or rules that will increase the reporting requirements and penalties.

RFG expects that the year 2025 will produce more cybersecurity issues, with the average financial impact of such attacks poised to increase from the previous year. On top of that, most digital users remain oblivious to the impacts that are caused by their poor business practices, including insecure email practices, insecure mobile apps, and insecure IoT (Internet of Things) embedded devices. In addition, CISOs need to add OT (operations technology) to their list of security priorities, as OT attacks will increase due to increased geopolitical strife.

It is evident from examining recent security breaches that many organizations have security lapses related to people and process failures (including the application of configurations, encryption, patches, and quality code). These factors continue to outweigh the impact of pure-play technology failures. To clean up their recent shortcomings and longer-term technical debt, CEOs and Boards of Directors must establish cyber-attack mitigation strategies, while CISOs, CSOs, and other security executives must re-imagine their security practices by analyzing them from all angles. The key takeaway is that companies need to adopt zero trust security and defense-in-depth strategies and thereby make their company "unappetizing" to attackers. In addition, organizations need to start investing in quantum-resistant cryptography, since current encryption methods will become obsolete as quantum computing systems mature and enter the marketplace. Real change will require a bottom-up and top-down approach, management and user training, and line-of-business conformity with required operating procedures. Moreover, C-level and Board members should directly participate in the oversight of strategic design and implementation, as well as acknowledge approval and participation via signatory sign-off. Anything less will simply aid in continuing the upward trajectory of cyber-attack successes.

Furthermore, the lack of cloud service provider (CSP) transparency remains an issue and an exposure in 2025. During the pandemic, many companies accelerated their move to the cloud – and instituted cloud-first policies for new applications under development. Part of the problem here is that the exact causes of some CSP outages are not publicly reported. This creates enough uncertainty that there is "room" for some providers to imply that some (if not most) breaches may have been caused by customers' failures, even if that is not the case. Cloud contracts are exceptionally vendor-biased to ensure that liabilities are reduced to minimums – that should give companies pause as responsibilities are virtually wholly shifted to user organizations in practice.

On the compliance front, countries and localities worldwide will continue to pass their own privacy acts, adding to a growing patchwork of requirements across geographies – and making full compliance a struggle for firms that operate in multiple regulatory regions. IT executives will be hard-pressed to keep up with compliance requirements globally until they can fully automate the processes, reducing the time it takes to align implementations to widely varying governance policies. To address these issues, IT executives will need to work with auditors and regulators to develop automated processes that will enable continuous compliance while reducing the cost of compliance and risk exposure. This will be an area where genAI can be of value. If that happens, 2025 will be the year when

enterprises start becoming audit-ready, allowing them to truly integrate compliance into the development cycle and production software.

DevSecOps / DataOps – The trend toward DevOps and DevSecOps – along with SRE (site reliability engineering) and DataOps for more effective data analytics – continues into 2025. However, progress will remain slow as most companies are still a long way from adopting these methodologies as standards for their application-development processes.

The challenge is not the technology but the organizational culture and adherence to legacy processes that is impairing customers' shift to DevSecOps. In general, developers resist picking up operations tasks while operations SRE staff seek to fill that void. The availability of staffing and skills will continue to be a challenge for IT executives throughout 2025. These constraints cause a redefinition of processes and tasks. As part of this analysis, compliance and security will have to become more integrated into the DevOps- DevSecOps processes – so that they will not become bottlenecks for new applications.

No-code/low-code development will gain more adherents and be adopted over time as new genAI tools come to market. In the past year genAI has demonstrated that it can be a coding productivity tool. RFG expects it to gain more traction in enterprises in 2025. The challenge to genAI code development adoption is ensuring governance, privacy and security measures are in place to protect against unexpected consequences.

DataOps, an automated, process-oriented methodology, is a needed counterpart to the more well-understood DevSecOps. RFG expects genAI to play a major role and help drive DataOps adoption as well. RFG is convinced that customers will come to see the benefits of using DataOps in coming years because of its potential to improve data quality and to reduce data-analytics cycle-times. We expect that DataOps will be more widely adopted by 2030.

DataOps, driven by genAI, gives IT organizations a new process to eliminate or minimize the islands of data (isolated data resources) in the enterprise data center – and in the clouds – that were created over time in previous development efforts. It is problematic to have new DevSecOps applications available for production within a few weeks or months, especially since some of the supporting data for those applications cannot be extracted, transformed, and loaded in less than six months.

RFG believes that interest in DataOps, especially ones utilizing genAI tools, will gain traction in 2025, leading to a year of piloting, proofs-of-concept (PoCs), and early trials for DataOps in many large enterprises.

Summary

2025 will be a transformative year for enterprises globally as they adapt to year of geopolitical uncertainty and conflicts both real and potential. It is important to recognize that business transformation goals are not a technology issue. Rather, the fundamental challenges are to address business-model issues, and these business transformation goals must be recognized as such. The mission of IT organizations is to support the shift; but business units, which provide the funding for new initiatives, are still driving the train.

C-suite and line of business executives must make the key business decisions that guide their company's business directions and digital transformations. As leaders, senior executives need to determine which businesses they plan to participate in and fund – and to make the appropriate structural changes to implement those plans. In 2025, business leaders need to develop strategies and tactical actions that lead to selecting which processes and applications will be designed and implemented across their entire organization. IT executives must ensure they align themselves with these new directions and directives across their enterprise.

RFG POV: The year 2025 will test business and IT executives' adaptability to change and shifting tides. IT organizations need to work harder to integrate their goals with those of the business – and to work collaboratively to enhance operations (whether on-site, in the cloud, or at the edge) and to innovate new, simpler approaches for doing business.

Although funding and other resources will be constrained in 2025, IT executives will need to invest significantly in AI and genAI solutions, cybersecurity, DevSecOps, DataOps, data center efficiency and sustainability, and other process improvements. These investments are needed to remain competitive in a rapidly changing world that is embracing genAI applications and tools. It will simultaneously help contain costs, mitigate the skills shortage, enhance compliance, increase flexibility and responsiveness, minimize risks, and improve resource utilization.

To achieve truly effective changes in business that will drive revenues and profits, business and IT executives must collaborate throughout the year. With this approach, IT budgets, plans, and strategies will dovetail with the primary goals of their organization – in business and in government – and ensure that they remain tightly integrated with the goals of their business throughout the year, regardless of external forces.

Additional relevant research is available at www.rfgonline.com. Interested readers should contact RFG Client Services to arrange further discussion or interview with Mr. Cal Braunstein.

46 Kent Hills Lane, Wilton, CT 06897; (203) 429-8951;
http://www.rfgonline.com/; Contact: inquiry@rfgonline.com