



Technology Brief...

October, 2017

J.Gold Associates LLC, 6 Valentine Road, Northborough, MA 01532, USA
www.jgoldassociates.com +1-508-393-5294
Research, Analysis, Strategy, Insight

Is the Smartphone the primary AI platform of the future?

INSIDE THIS ISSUE

- 1** Is the Smartphone the primary AI platform of the future?
- 2** Can Ryzen Pro propel AMD into a serious enterprise client challenger?
- 3** SAP Leonardo – From IoT to mainstream

Artificial Intelligence (AI) is emerging as a key technology to assist both workers and consumers in everyday activities. Indeed, some tools, like Google Assistant, Amazon's Alexa, Samsung's Bixby, etc. already have limited AI capability. But given that most people currently already carry a computer in their pockets (e.g., a Smartphone), and its where many interface most often to the wider world, can smartphones become the primary platform for AI?

Not in the short term. It will certainly morph into a primary user interface for actionable intelligence that makes better user interactions possible. But while there will be an increasing level of AI capability built into the device as we progress in computing ability in general as a result of Moore's Law, including specialized processing subsystems optimized for AI features, it's also true that the amount and types of AI processes and algorithms we'll need over time will grow in complexity and require far more processing capability than built into the devices. That means we'll need increasing levels of edge computing built into servers, some of them highly specialized, at the edge of the network. But we'll also need some pretty big cloud computing facilities to process the truly massive AI data sets for things like autonomous driving, smart cities, medical/health discoveries, etc.

"...In the next 1-2 years we'll see SoCs from Qualcomm targeted at mobile devices and Intel targeted at edge servers, among others... This will herald a new generation of products that are much more cost effective and able to function much better in a limited power-available product like a smartphone or other autonomous devices. At that point, smartphones will likely be the most predominant AI powered devices....."

But longer term things may change. Some current implementations of AI primarily relegated to large computing resources are evolving. AI is moving from the cloud and server down to taking place on the end user device as more powerful chips create a new platform that is high enough performance for image classification, voice, security, etc. On-device significantly reduces the delays and latency of interactions. But to achieve this, requires lots of processing power, including specialized chips/components, and large amounts of memory. The biggest obstacle is in the requirements for training which requires massive amounts of compute cycles and is typically done in a server farm somewhere, with learned algorithms then downloaded to the device. But with rapid advances in dedicated functions in silicon, we'll soon have the able to do at least some level of training on the devices itself and not just having devices running the learned algorithms. We expect to see this capability emerge in portable devices within the next 3-4 years and will be an important step in democratizing AI usage. And with the general availability of 5G, also in the next 3-4 years, we don't need to leave the cloud fully behind, but can provide balance to fully leverage what happens in the cloud vs. on the device.

Chip companies are investing heavily in AI technology. Qualcomm recently

purchased Scyfer, and has partnered extensively with Google on TensorFlow and Facebook on Caffe. And others like Intel with its acquisitions of Nervana, Movidius, Safron, Mobileye and Altera are also investing heavily, although in the case of Intel much more targeted towards the edge to back end (cloud), than is Qualcomm with its heavy focus on the front end to edge. Google and Microsoft have also announced specialized AI silicon targeted at their cloud infrastructure.

But portability will ultimately be crucial for widespread adoption. Currently Qualcomm has a Neural Processing engine that it uses to help developers port AI apps and that will eventually be fully compatible with on-chip implementations of the identical SDK. This strategy is critical, as those companies building developer momentum now often before fully optimized processors are available, are the ones that will ultimately gain market share. However, algorithmic advances are still needed to make this work. Qualcomm for one has doing research in compression, optimization in training networks, etc. to see what can be done in the mobile platform. But it also has to partner with research and other companies to migrate some of the major products to run in a mobile platform environment, while also building bridges with training frameworks like TensorFlow, Caffe, etc. This is no easy task.

To make the on-board devices generally available and as capable as possible, several things must happen. First, there needs to be much more work done on creating neural processing units that can potentially do some amount of training on board to off-load what's needed in the cloud, but more importantly to efficiently run the trained learning system that is key to achieving AI on the device, without the latency of having to run everything in the cloud. Even though fast networks like 5G will greatly lesson the latency, any latency could be catastrophic for personal safety or device autonomous accident avoidance systems. We've grown accustomed to seeing a CPU, GPU and often a DSP along with other similar specialized functional blocks inside the common SoC. What we'll be seeing much more often in the near future will be NPU (Neural Processing Units) that over time will get increasingly sophisticated and eventually directly run some of the machine learning frameworks (e.g., TensorFlow, Caffe).

Bottom Line: Currently, much of the processing for AI/ML is done on specialized chips (e.g., Nvidia) that are peripheral to the overall SoC. In the next 1-2 years we'll see SoCs from Qualcomm targeted at mobile devices and Intel targeted at edge servers, among others. They will incorporate the specialized subsystems internally rather than in a peripheral chip. This will herald a new generation of products that are much more cost effective and able to function much better in a limited power-available product like a smartphone or other autonomous devices. At that point, smartphones will likely be the most predominant AI powered devices used in both consumer and business environments.

“...Ryzen Pro definitely raises the bar for AMD and creates an alternative to Intel chips in business clients. Indeed, these chips are probably the most credible threat AMD has presented in more than 5 years. But unless AMD can show a clear advantage beyond just a price advantage, it may be a tough sell. ...”

Can Ryzen Pro propel AMD into a serious enterprise client challenger?

AMD recently launched its Ryzen Pro client platform targeted specifically at enterprise, to try and be more competitive with Intel enterprise oriented client devices and to allow it capture some of the dominant Intel market share. Ryzen Pro has added some features that are important to corporate client purchasers, e.g., hardware enabled encryption, secure booting, TPM for security, image stability for 18 months providing less software image “breakage”, longer confirmed availability of 24 months allowing products to be available for a more enterprise friendly lifecycle, etc.. But is this enough for enterprises to adopt Ryzen Pro?

Dell, HP, and Lenovo all announced they would be selling clients in their business lines with AMD Ryzen Pro processors. The question is, will they be able to compete with Intel vPro products that are targeted at the enterprise user? While not all enterprise clients are purchased with vPro, it does provide Intel with a significant competitive advantage from both a marketing and device management perspective. vPro is only activated on 5%-10% of business clients sold, although I estimate that at least twice that many are sold with vPro chips and never activated. vPro has a long history of securing corporate devices, even though it may have started life more as a management function. But with increasing concern for malware attacks, and particularly ransomware, many more companies are seriously looking at vPro, especially since Intel has been pursuing much more of a security approach to the product over recent years.

But the availability of Ryzen Pro does provide PC OEMs with some advantages. If they can offer a credible alternative to Intel in their machines, they can lower prices for those devices (AMD still sells its products not on premium pricing as Intel does, but at a discount to equivalent Intel products). Further, if Ryzen Pro does take off, the vendors will have a way to pressure Intel for more favorable pricing and/or concessions. However AMD has not spent much effort in the past in heavily promoting its product features, unlike Intel. AMD promises to do so this time around, but it's not yet clear how much effort they will make and how successful they will be.

Bottom Line: Ryzen Pro definitely raises the bar for AMD and creates an alternative to Intel chips in business clients. Indeed, these chips are probably the most credible threat AMD has presented in more than 5 years. But unless AMD can show a clear advantage beyond just a price advantage, it may be a tough sell. While enterprises are certainly focused on keeping prices low, they are much less sensitive to price than the competitive pressures on consumer purchases. Further, With Intel's advantage in not only vPro on chip, but all of the software support it provides and gets from third parties, AMD will have a lot of catching up to do. I believe companies should evaluate the Ryzen Pro alternatives, particularly if they are not fully utilizing all of the management features of the vPro line. But companies should also look at more than just a lower price before making a decision.

SAP Leonardo – From IoT to mainstream

SAP recently announced it is expanding its Leonardo brand to encompass many more functions and capabilities. SAP Leonardo started life as an "Accelerator that was targeted at specific verticals, and originally targeted the emerging needs of the Enterprise of Things (EoT). It was a way for SAP customers, through packaged products and services, to get to market more quickly with enterprise focused IoT implementations.

While not yet a massive market, the Leonardo strategy for EoT proved successful. As a result, SAP has recently rebranded Leonardo, and it's now a general purpose "market basket" of products and services meant to quickly develop and deploy any of the needed functions available within the SAP ecosystem and extensible by others. This approach has a good deal of merit, particularly in the smaller to medium size deployments where the expertise and learning from peer deployments can be exploited. And as a way to accelerate the needed Digital Transformation process that most companies are now experiencing, SAP can claim a faster, easier way to get there. Indeed, we estimate less than 10% of organizations have the required expertise fully in house to digitally transform their

"...SAP is positioning Leonardo as its way to get organizations to market faster and with "built-in" expertise. This is a welcome response to the past challenges when SAP was seen as a difficult product to implement and a long term deployment operation....."

Recent Research

Contact us to request the following research reports:

Market Studies

- **The State of Enterprise Mobile Management (EMM)**
- **Mobile E-Commerce: Friend or Foe?**

Emerging Technology Trends

- Highlights our key emerging trends for the next 3-4 years

Commentary and Analysis

- **Apple and IBM in Enterprise: Joined at the Apps**

Research Reports

- **Android in the Business Environment: Is it Safe?**
- **Your PC has an Identity Crisis: Saving the cost of hacks and other benefits of enhanced identity**
- **Replacing Enterprise PCs: The Fallacy of the 3-4 Year Upgrade Cycle**
- **Keeping Notebooks Past Their Prime: A Study of Failures and Costs**

Whitepapers

- **A Heuristic Approach to Mobile Security**
- **MDM- Where Do We Go From Here?**



J. Gold Associates, LLC
6 Valentine Road
Northborough, MA 01532 USA

Phone:
+1-508-393-5294

Web:
www.jgoldassociates.com

**Research, Analysis,
Strategy, Insight**

business. The rest will need assistance.

Leonardo will consist of several specialty solutions and components, like IoT, Machine Learning, cognitive capability, etc., but it will be an open system that allows add-ons and third party solutions to be integrated, as well as extensible with a company's own custom built applications. Leonardo will follow the same basic model of getting customers up and running more quickly than they could on their own and hence quickly get a return on investment and achieve a competitive business advantage. Indeed, SAP claims it can reduce the transformation process dramatically in a fixed price, fixed time frame delivery mechanism, and want to get to value within 8 weeks, although this may not be a completed full scale solution with all of the desired processes included.

Leonardo makes required expertise accessible to many businesses in the way of templates and deployment specialists that many organizations would not be able to achieve on their own. This leveragability is a considerable benefit when speed of deployment in a hyper competitive business climate is critical (taking 2-3 years to get anything deployed today is a non starter). Leonardo can often reduce time to full deployment to just 3-6 months, depending on the complexity and whether its standard practice across an industry. Given that we estimate most organizations take 12-24 months to get to even a partially digitally transformed state, this provides a great advantage if it can be achieved.

Bottom Line: SAP is positioning Leonardo as its way to get organizations to market faster and with "built-in" expertise. This is a welcome response to the past challenges when SAP was seen as a difficult product to implement and a long term deployment operation. With its specialized vertical deployments, and its template approach, Leonardo will expand the market for SAP and allow companies to get up and running in a shortened timeframe. This is a major advance for SAP and will expand its market opportunity. Companies looking at specific verticals should evaluate the SAP Leonardo offerings.

About J. Gold Associates, LLC.

J. Gold Associates provides advisory services, syndicated research, strategic consulting and in-context analysis to help its clients make important technology choices and to enable improved product deployment decisions and go to market strategies. We work with our clients to produce successful new product strategies and deployments through workshops and reviews, business and strategic plan coaching and reviews, assistance in product selection and vendor evaluations, needs analysis, competitive analysis, and ongoing expertise transfer.

J. Gold Associates provides its clients with insightful, meaningful and actionable analysis of trends in the computer and technology industries. We have acquired a broad based knowledge of the technology landscape and business deployment requirements, and bring that expertise to bear in our work. We cover the needs of business users in enterprise and SMB markets, plus focus on emerging consumer technologies that will quickly be re-purposed to business use.

We can provide your company with a trusted and expert resource to maximize your investments and minimize your risk. Please contact us to see how we can help you.